

Precision medical imaging for cancer diagnosis and treatment

Edited by

Alla Reznik, Aaron Fenster, Laura Curiel,
Ira Blevis and Mitchell Albert

Published in

Frontiers in Oncology



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-3946-0
DOI 10.3389/978-2-8325-3946-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Precision medical imaging for cancer diagnosis and treatment

Topic editors

Alla Reznik — Lakehead University, Canada
Aaron Fenster — Western University, Canada
Laura Curiel — University of Calgary, Canada
Ira Blevis — Philips Healthcare, United States
Mitchell Albert — Lakehead University, Canada

Citation

Reznik, A., Fenster, A., Curiel, L., Blevis, I., Albert, M., eds. (2023). *Precision medical imaging for cancer diagnosis and treatment*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3946-0

Table of contents

- 06 **¹⁸F-DCFPyL PET/CT in advanced high-grade epithelial ovarian cancer: A prospective pilot study**
Ur Metser, Roshini Kulanthaivelu, Tanya Chawla, Sarah Johnson, Lisa Avery, Douglas Hussey, Patrick Veit-Haibach, Marcus Bernardini and Liat Hogen
- 15 **Breast cancer diagnosis using frequency decomposition of surface motion of actuated breast tissue**
Jessica Fitzjohn, Cong Zhou and J. Geoffrey Chase
- 34 **Are contrast enhanced mammography and digital breast tomosynthesis equally effective in diagnosing patients recalled from breast cancer screening?**
Natalia Siminiak, Anna Pasiuk-Czepczyńska, Antonina Godlewska, Piotr Wojtyś, Magdalena Olejnik, Joanna Michalak, Piotr Nowaczyk, Paweł Gajdzis, Dariusz Godlewski, Marek Ruchata and Rafał Czepczyński
- 45 **Predicting incomplete cytoreduction in patients with advanced ovarian cancer**
Eva K. Egger, Marie Antonia Buchen, Florian Recker, Matthias B. Stope, Holger Strunk, Alexander Mustea and Milka Marinova
- 55 **Advantages of FBPA PET in evaluating early response of anti-PD-1 immunotherapy in B16F10 melanoma-bearing mice: Comparison to FDG PET**
Mitsuaki Tatsumi, Fumihiko Soeda, Sadahiro Naka, Kenta Kurimoto, Kazuhiro Ooe, Hideyuki Fukui, Daisuke Katayama, Tadashi Watabe, Hiroki Kato and Noriyuki Tomiyama
- 64 **High-resolution synthesis of high-density breast mammograms: Application to improved fairness in deep learning based mass detection**
Lidia Garrucho, Kaisar Kushibar, Richard Osuala, Oliver Diaz, Alessandro Catanese, Javier del Riego, Maciej Bobowicz, Fredrik Strand, Laura Igual and Karim Lekadir
- 81 **When tissue is not the only issue: Poorly differentiated lung squamous-cell carcinoma with adrenal, costochondral, and cardiac metastases – case report**
Megan Clark, Andres G. Griborio-Guzman, Nishigandha P. Burute, Sonja Lubbers, Margaret L. Anthes, Masoud Sadreddini and Olexiy I. Aseyev
- 88 **Clinical value of ¹⁸F-FDG PET/CT in the management of HIV-associated lymphoma**
Qi Liu, Tao Yang, Xiaoliang Chen and Yao Liu
- 97 **Clinical advancement of precision theranostics in prostate cancer**
Yue Tan, Zhihui Fang, Yongxiang Tang, Kai Liu and Hong Zhao

- 110 **Small hyperattenuating adrenal nodules in patients with lung cancer: Differentiation of metastases from adenomas on biphasic contrast-enhanced computed tomography**
Lixiu Cao, Libo Zhang and Wengui Xu
- 122 **Preoperative prediction of tumor deposits in rectal cancer with clinical-magnetic resonance deep learning-based radiomic models**
Chunlong Fu, Tingting Shao, Min Hou, Jiali Qu, Ping Li, Zebin Yang, Kangfei Shan, Meikang Wu, Weida Li, Xuan Wang, Jingfeng Zhang, Fanghong Luo, Long Zhou, Jihong Sun and Fenhua Zhao
- 131 **The value of contrast-enhanced ultrasonography in differential diagnosis of primary testicular germ cell tumors and non-germ cell tumors over 50 years old**
Nianyu Xue, Guoyao Wang, Shengmin Zhang and Yijun Lu
- 138 **The road to breast cancer screening with diffusion MRI**
Mami Iima and Denis Le Bihan
- 153 **Impact of localized fine tuning in the performance of segmentation and classification of lung nodules from computed tomography scans using deep learning**
Jingwei Cai, Lin Guo, Litong Zhu, Li Xia, Lingjun Qian, Yuan-Ming Fleming Lure and Xiaoping Yin
- 164 **Development and validation of a nomogram model for lung cancer based on radiomics artificial intelligence score and clinical blood test data**
Wenteng Hu, Xu Zhang, Ali Saber, Qianqian Cai, Min Wei, Mingyuan Wang, Zijian Da, Biao Han, Wenbo Meng and Xun Li
- 173 **Identification of a novel peptide targeting TIGIT to evaluate immunomodulation of ^{125}I seed brachytherapy in HCC by near-infrared fluorescence**
Peng Zeng, Duo Shen, Wenbin Shu, Shudan Min, Min Shu, Xijuan Yao, Yong Wang and Rong Chen
- 185 **Clinical radiomics-based machine learning versus three-dimension convolutional neural network analysis for differentiation of thymic epithelial tumors from other prevascular mediastinal tumors on chest computed tomography scan**
Chao-Chun Chang, En-Kuei Tang, Yu-Feng Wei, Chia-Ying Lin, Fu-Zong Wu, Ming-Ting Wu, Yi-Sheng Liu, Yi-Ting Yen, Mi-Chia Ma and Yau-Lin Tseng
- 197 **Differential regional importance mapping for thyroid nodule malignancy prediction with potential to improve needle aspiration biopsy sampling reliability**
Liping Wang, Yuan Wang, Wenliang Lu, Dong Xu, Jincan Yao, Lijing Wang and Lei Xu

- 210 **Clinical validation of the novel CLIA-CA-62 assay efficacy for early-stage breast cancer detection**
Marina Sekacheva, Alexander Boroda, Anastasia Fatyanova,
Alexander Rozhkov and Nikolai Bagmet
- 221 **Value of dual-layer spectral detector computed tomography in the diagnosis of benign/malignant solid solitary pulmonary nodules and establishment of a prediction model**
Xiaodong Xie, Kaifang Liu, Kai Luo, Youtao Xu, Lei Zhang,
Meiqin Wang, Wenrong Shen and Zhengyang Zhou



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

Oleksandr Bubon,
Thunder Bay Regional Research
Institute, Canada
Olexiy Aseyev,
Thunder Bay Regional Health Sciences
Centre, Canada

*CORRESPONDENCE

Ur Metser
ur.metser@uhn.ca

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 22 August 2022

ACCEPTED 26 September 2022

PUBLISHED 13 October 2022

CITATION

Metser U, Kulanthaivelu R, Chawla T,
Johnson S, Avery L, Hussey D,
Veit-Haibach P, Bernardini M and
Hogen L (2022) ^{18}F -DCFPyL PET/CT
in advanced high-grade epithelial
ovarian cancer: A prospective pilot
study.
Front. Oncol. 12:1025475.
doi: 10.3389/fonc.2022.1025475

COPYRIGHT

© 2022 Metser, Kulanthaivelu, Chawla,
Johnson, Avery, Hussey, Veit-Haibach,
Bernardini and Hogen. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

^{18}F -DCFPyL PET/CT in advanced high-grade epithelial ovarian cancer: A prospective pilot study

Ur Metser ^{1*}, Roshini Kulanthaivelu¹, Tanya Chawla¹,
Sarah Johnson¹, Lisa Avery², Douglas Hussey¹,
Patrick Veit-Haibach¹, Marcus Bernardini³ and Liat Hogen³

¹Joint Department of Medical Imaging, University Health Network, Sinai Health Systems, Women's College Hospital, University of Toronto, Toronto, ON, Canada, ²Department of Biostatistics, University Health Network, Toronto, ON, Canada, ³Division of Gynecologic Oncology, University Health Network, University of Toronto, Toronto, ON, Canada

Objectives: Glutamate carboxypeptidase-II (GCP-II), a zinc metalloenzyme that resides in cell membrane, has been reported as overexpressed in the neovasculature of ovarian cancers. The study objective was to determine whether GCP-II targeted imaging with ^{18}F -DCFPyL PET/CT can detect disease sites in women with advanced high-grade serous ovarian cancer (HGSOC).

Materials and methods: Twenty treatment-naïve women with advanced HGSOC were recruited (median age 60 years). Prior to commencing therapy (primary cytoreductive surgery [n=9] or neoadjuvant chemotherapy [n=11]), subjects underwent routine staging with contrast-enhanced abdominopelvic CT (=CT), followed by ^{18}F -DCFPyL PET/CT (=PET). CT and PET were reported independently using a standardized reporting template assessing 25 sites. The performance of PET was compared to CT in all subjects and to surgery and surgical histopathology in 9 patients who underwent primary cytoreductive surgery.

Results: Of the 25 sites assessed in 20 patients, CT detected disease in 292/500 (58.4%) locations and PET detected disease in 171/500 (34.2%). Compared to CT the sensitivity (95% CI) of PET to detect disease in the upper abdomen, the gastrointestinal tract or the peritoneum was 0.29 (0.20,0.40), 0.21 (0.11,0.33) and 0.74 (0.64,0.82), respectively. In the surgical cohort, 220 sites in 9 patients were evaluated. The sensitivity and specificity of CT and PET were 0.85 versus 0.54 (p<0.001) and 0.73 versus 0.93 (p<0.001), respectively.

Conclusion: Although ^{18}F -DCFPyL has higher specificity than CT in detecting advanced HGSOC tumor sites, it detects less disease sites than CT, especially in the upper abdomen and along the gastrointestinal tract, likely limiting its clinical utility.

Clinical trial registration: [ClinicalTrials.gov](#), NCT03811899.

KEYWORDS

^{18}F -DCFPyL, PSMA, PET/CT, CT, ovarian cancer, diagnostic performance

Introduction

Epithelial ovarian cancer is the fifth most common cancer in women in the United States, and it has the highest fatality-to-case ratio of all the gynecologic malignancies (1). These tumors spread primarily by exfoliation of cells into the peritoneal cavity, but also by lymphatic and hematogenous dissemination. High-grade serous carcinomas (HGSOC) are the most common histologic subtype, with up to 75% of patients presenting with advanced-stage disease, for which surgery alone is not curative. Standard therapy consists of either primary cytoreductive surgery (PCS) followed by platinum-based chemotherapy or neoadjuvant chemotherapy (NACT) followed by interval cytoreductive surgery and further chemotherapy. Achieving complete resection in advanced ovarian cancer is often not feasible due to the multi-focal, disseminated nature of the disease (2). Numerous studies have shown that the degree of cytoreduction (i.e., the amount of residual disease at the completion of surgery) is directly correlated with survival. Patients with absence of gross residual disease after surgery have a much better outcome than those with optimal debulking (defined as residual sites of disease < 1 cm in diameter) or suboptimal debulking (defined as residual sites of disease \geq 1 cm), with a 5-year survival estimated at only 15% after suboptimal debulking (2–5).

Accurate mapping of the distribution and volume of metastatic disease is vital for determination of the optimal therapeutic approach (PCS vs. NACT). Currently, most patients are staged with contrast-enhanced CT of the chest, abdomen, and pelvis (=CT); however, this tool has limited sensitivity and specificity, especially for disease in the mesentery or serosal surface of bowel (6–8). FDG PET/CT has been previously assessed with sensitivity and specificity of 78% and 68%, respectively, on a quadrant basis. Given the moderate performance measures, FDG PET has not been universally incorporated into the workup of these patients (9–11). A further non-invasive tool that would accurately map disease extent is needed to better select patients for primary therapy, reduce the rate of aborted surgery and associated morbidity, and hopefully improve patient outcomes.

Glutamate carboxypeptidase II (GCP-II) is a zinc metalloenzyme that resides in cell membranes, mostly on the extracellular side. It has various additional names including folate hydrolase and prostate specific membrane antigen (PSMA). It is expressed by normal tissues such as salivary and lacrimal glands, larynx, kidneys, bowel and prostate, as well as by multiple malignant tumors, often in the neovasculature of these tumors. GCP-II (=PSMA) has been extensively assessed in the setting of prostate cancer, especially in the setting of biochemical recurrence. In prostate cancer, PSMA PET has shown a very

high sensitivity and moderately high specificity for the detection of recurrent or metastatic disease even when conventional imaging is negative (12–22). Initial report on the expression of GCP-II in neovasculature of gynecologic cancers including primary and metastatic epithelial ovarian cancer suggested high expression of GCP-II at immunohistochemistry in all 46 cases of ovarian cancer assessed (23). These findings were the impetus for the current study. The main aim of the current study was to determine whether GCP-II targeted imaging with ^{18}F -DCFPyL PET/CT (=PET) can detect sites of disease in women with advanced HGSOC and to compare sites of disease detected on PET to CT and to intra-operative findings and surgical histopathology.

Patients and methods

Study design

This is an institutional ethics review board approved, single arm, prospective pilot study (ClinicalTrials.gov: NCT03811899). Written informed consent was obtained from all participants. The inclusion criteria were: 1. Age \geq 18 years; 2. Cytological or histological diagnosis of high grade epithelial ovarian cancer; 3. Clinical stage III or IV, under consideration for PCS or NACT; 4. Contrast-enhanced CT abdomen and pelvis within 6 weeks of PET. Exclusion criteria included: 1. Evidence of epithelial ovarian cancer of the following histological subtypes: mucinous, low grade serous, low grade endometrioid and low-malignant potential tumors or metastases from other primary tumor; 2. Inability to complete study procedures (contraindication for PET as per institutional guidelines such as pregnancy, or participant's inability to lie still for 30 minutes). Demographic and clinical data include age, FIGO (International Federation of Gynecology and Obstetrics) stage, serum CA-125 at presentation, and surgical outcomes were tabulated.

Study procedures

^{18}F -DCFPyL PET/CT

PET was performed 90–120 minutes (mean \pm SD: 100.3 \pm 9.7) after injection of 310 (\pm 16.8) MBq of ^{18}F -DCFPyL. During uptake time, water soluble oral contrast was given for bowel opacification on CT. Patients were positioned supine on the imaging couch with arms outside of the region of interest. Images were obtained from the skull base to the upper thighs. PET was performed on a Biograph mCT 40 scanner (Siemens Healthcare, Erlangen, Germany). Low dose CT without intravenous contrast was used for attenuation correction as per standard departmental protocols. Overall, 5–9 bed positions were obtained as per patient height (2–5 min/bed position).

Abbreviations: Gcp-ii, Glutamate carboxypeptidase II; PCS, primary cytoreductive surgery; NACT, neoadjuvant chemotherapy.

CT protocol

The contrast-enhanced CT scan were performed by using the Aquilion 64 or Aquilion ONE CT (Canon Medical Systems). The scanning parameters were: tube voltage 120 kV and tube determined using automatic exposure control (^{SURE}exposure). For Aquilion ONE scan parameters were as follows: 1–3 mm slice thickness; 2.4 mm slice interval; helical pitch = 65; pitch factor = 0.813. For Aquilion 64 scan parameters were as follows: 1–5 mm slice thickness, 2.5 mm slice interval; helical pitch: 53; pitch factor = 0.828. Images were obtained after intravenous administration of 100 ml of 300 mg of iodine per milliliter of nonionic contrast material (Ultravist 370; Schering) using a power injector through an 18-gauge at a rate of 3 ml/s. Coronal and sagittal reformats of the dataset were also obtained.

Imaging interpretation & reporting template

CT and PET imaging data sets were interpreted independently. When present, primary tumor, nodal, peritoneal and visceral metastases on CT were recorded by one of 2 readers (TC, SJ; with 21 and 9 years of experience) using standard diagnostic criteria (24). PET was interpreted in consensus by 2 readers (UM, RK with 20 and 5 years of experience). In general, on PET, focal tracer accumulation greater than background activity, which could not be attributed to physiological activity, or a benign entity were recorded. SUVmax at all tumor sites and PSMA score relative to reference tissues, as previously described, were documented (25). All disease sites on either modality were tabulated using a standardized synoptic reporting template evaluating 25 stations in the abdomen and pelvis (Appendix A). These included assessment of: 1. Primary ovarian tumor/s; 2. Nodal metastases (below and/or above the renal veins); 3. Peritoneum (8 stations); 4. Gastrointestinal tract (5 stations); 5. Upper abdomen (9 stations).

Reference standard

A head-to-head comparison of lesion detection on CT and PET was performed for all patients. For the subset of patients who underwent primary cytoreductive surgery findings on CT and PET were compared to intra-operative findings and surgical histopathology using the same synoptic reporting template. The detection rate, sensitivity, specificity, positive predictive value, negative predictive value, and overall accuracy were calculated for each modality according to the standard of reference for all evaluable stations.

Statistical analysis

Sensitivity and specificity of PET were calculated on a per-lesion basis against standard of care contrast-enhanced CT. Calculations were performed across all lesions as well as at the local and regional level with exact confidence intervals as per Collett (26). A per-lesion analysis assumes lesions within patients are independent. To adjust for potential similarities of

the assessment of lesions within patients a second analysis of the sensitivity and specificity was undertaken using generalised estimating equations (GEE) and an exchangeable working correlation structure as in Smith and Hadgu (26).

A sub-sample of nine patients underwent PCS with surgical histopathology. Sensitivity and specificity of both CT and PET were evaluated against the surgical standard for this subsample. Exact binomial tests (27, 28) were used to test the null hypothesis that the sensitivity and specificity of the tests was equal across regions, but the subsample was too small to detect differences at the lesion level.

Results

Of the 112 participants approached to participate, 92 were excluded (Figure 1). There were 20 women (median age, 60 years; range: 38–83) with histologically proven high grade epithelial ovarian cancer included of whom 11 had NACT and 9 underwent PCS. Clinical data including stage, serum CA-125 and surgical outcome are summarized in Table 1.

Detection rate on PET

Of the 500 stations assessed in all 20 patients, CT detected disease in 292/500 (58.4%) and PET detected disease in 171/500 (34.2%). ¹⁸F-DCFPyL uptake when visible at disease sites was generally low or low to moderate with a mean SUVmax (\pm SD) of 4.2 ± 1.9 (range: 1.2–10.9). Of all 171 lesions assessed on PET, PSMA scores were 0, 1, and 2 in 8/171 (4.7%), 130/171 (76%), 33 (19.3%), respectively. No lesion with PSMA score of 3 was recorded.

Primary tumors were detected in 19/20 participants (95%) on both PET and CT. The performance measures of PET compared to CT for detection of disease for all evaluated stations and for metastatic sites grouped by anatomical location are presented in Table 2 including the GEE adjusted measures, correcting for lack of independency of multiple lesions within the same patient.

Performance of CT and PET with surgery as reference standard

Nine participants underwent PCS with 220 evaluable stations (data were missing for 5 stations in one patient). There were 54 stations that were positive on CT, PET and at surgery; and 76 stations that were negative on CT, PET and at surgery. The overall performance measures of CT, and PET with surgery and surgical histopathology as the reference standard are presented in Table 3. Disease detection on imaging in the various stations evaluated with comparison to surgery as the reference standard is depicted in Figure 2; the sensitivity and specificity of CT and PET in identifying metastatic sites compared to surgery and surgical histopathology are presented in Table 4.

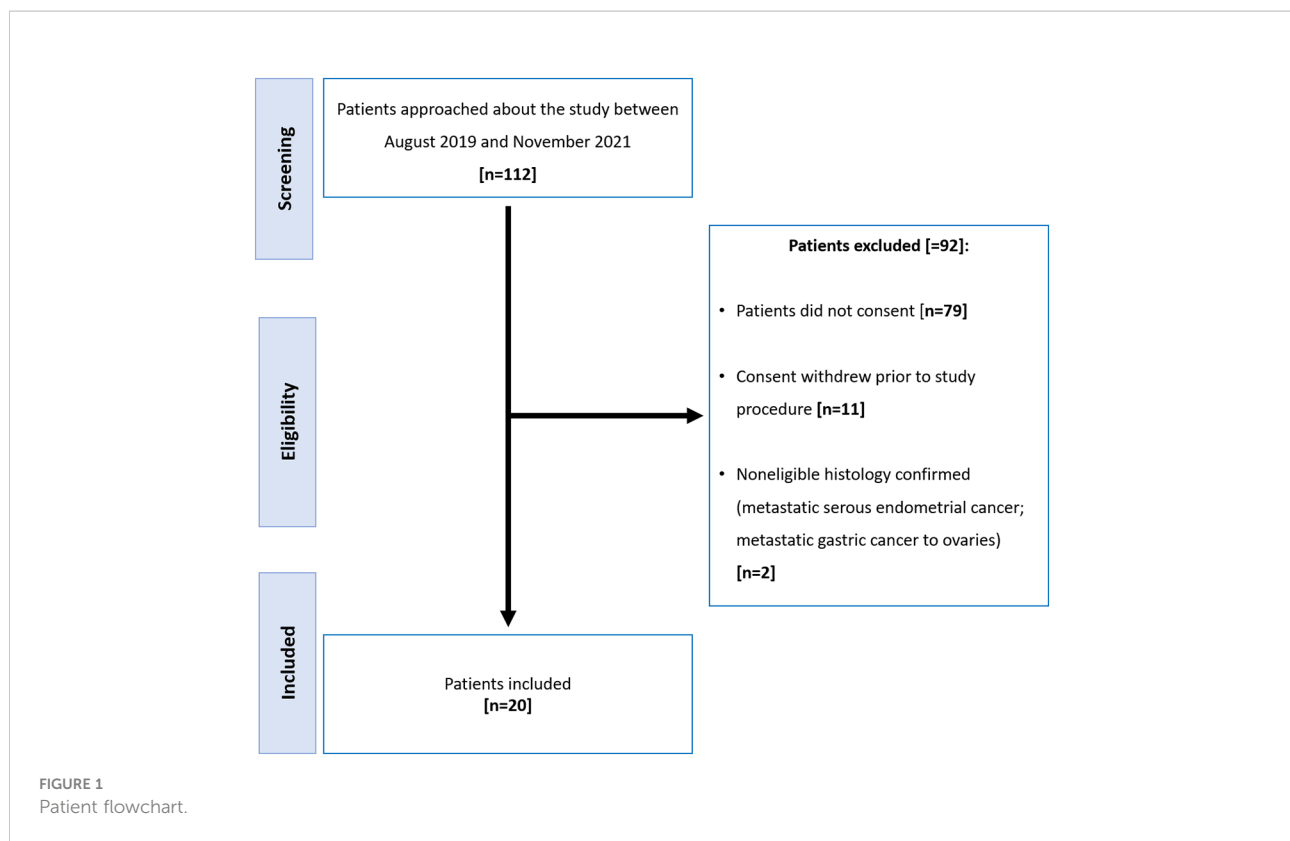


TABLE 1 Summary of clinical parameters.

		Number of participants=
Stage	IIIC	13
	IV	7
Serum CA-125 (U/ml)	Mean ± SD	2238± 4100
	Range (min., max)	(64, 19154)
Surgical Outcome	Complete cytoreduction	8
	Optimal debulking (< 1cm)	10
	Suboptimal debulking	2

Discussion

In women with advanced HGSOc, ^{18}F -DCFPyL (PSMA) PET/CT detects fewer metastatic sites of disease as compared to standard of care contrast-enhanced CT, but at a higher specificity. Although only a subsample of the study population underwent PCS, in these participants, surgery and surgical histopathology was used as the reference standard to compare the performance of CT and PET. Comparison of the sensitivity and specificity of modalities in this subsample is based on the assumption of independent lesions. The comparison of per lesion and GEE analyses indicates that this assumption is reasonable for most

regions, with little difference between the adjusted and per lesion values. The performance of PET was especially poor for lesions in the upper abdomen and along the gastrointestinal tract. This is likely due to limited expression of GCP-II at tumor sites, as depicted with ^{18}F -DCFPyL PET/CT, along with the high background activity in the liver and spleen and in segments of the gastrointestinal tract, limiting detection of subdiaphragmatic or capsular hepatic metastases and serosal deposits (Figures 3A–G). These findings suggest that GCP-II targeted imaging with ^{18}F -DCFPyL PET/CT in women with advanced HGSOc is likely of limited clinical utility.

Preclinical studies on the expression of GCP-II in ovarian cancer have shown conflicting results. Wernicke et al. examined

TABLE 2 Performance measures of PET with CT as the reference for all stations assessed, and for the various stations grouped by anatomic location (excluding primary tumors).

STATISTIC	ALL STATIONS (n=500)	LYMPH NODES (n=40)	UPPER ABDOMEN (n=180)	GI TRACT (n=100)	PERITONEUM (n=160)
<i>Value(95% CI)</i>					
Prevalence	0.58 (0.54, 0.63)	0.40 (0.25, 0.57)	0.51 (0.44, 0.59)	0.63 (0.53, 0.72)	0.64 (0.56, 0.71)
<i>GEE adjusted</i>	0.58	0.40	0.51	0.63	0.64
Sensitivity	0.49 (0.43, 0.55)	0.56 (0.30, 0.80)	0.29 (0.20, 0.40)	0.21 (0.11, 0.33)	0.74 (0.64, 0.82)
<i>GEE adjusted</i>	0.49 (0.42, 0.56)	0.44 (0.21, 0.66)	0.30 (0.18, 0.42)	0.21 (0.10, 0.31)	0.74 (0.62, 0.86)
Specificity	0.87 (0.81, 0.91)	0.79 (0.58, 0.93)	0.93 (0.86, 0.97)	0.97 (0.86, 1.00)	0.72 (0.59, 0.83)
<i>GEE adjusted</i>	0.86 (0.80, 0.93)	0.72 (0.53, 0.91)	0.93 (0.88, 0.99)	0.97 (0.93, 1.02)	0.73 (0.60, 0.86)
PPV	0.84 (0.77, 0.89)	0.64 (0.35, 0.87)	0.82 (0.65, 0.93)	0.93 (0.66, 1.00)	0.82 (0.73, 0.90)
NPV	0.55 (0.49, 0.60)	0.73 (0.52, 0.88)	0.56 (0.47, 0.64)	0.42 (0.31, 0.53)	0.61 (0.48, 0.72)
Accuracy	0.65 (0.60, 0.69)	0.70 (0.53, 0.83)	0.61 (0.53, 0.68)	0.49 (0.39, 0.59)	0.73 (0.66, 0.80)

GI, gastrointestinal; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value. GEE, generalised estimating equations.

TABLE 3 Performance measures of CT, and PET with surgery and surgical histopathology as the reference standard for all evaluated stations (n=220).

STATISTIC	CT (95% CI)	PET (95% CI)	<i>p-value</i>
Sensitivity	0.85 (0.77, 0.91)	0.54 (0.44, 0.63)	<0.001
Specificity	0.73 (0.63, 0.81)	0.93 (0.86, 0.97)	<0.001
PPV	0.77 (0.68, 0.84)	0.88 (0.78, 0.95)	
NPV	0.82 (0.73, 0.89)	0.66 (0.57, 0.73)	
Accuracy	0.79 (0.73, 0.84)	0.73 (0.66, 0.78)	

PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval.

the expression of PSMA in neovasculature of gynecologic cancers including primary and metastatic ovarian cancer (23). The authors showed a high expression of PSMA at immunohistochemistry in all 46 cases of ovarian cancer assessed, a report which provided the impetus for the current study. A further, more recent study published by Aide et al. assessed 32 patients with 57 samples (including 25 samples obtained after chemotherapy). The authors demonstrated the quasi-absence of PSMA expression within serous epithelial ovarian cancers. Authors showed no correlation with resistance to chemotherapy and non-evolution of PSMA expression during the treatment course (29). Our findings are more in line with the results of Aide et al, with most disease sites in women with HGSOC showing low level ¹⁸F-DCFPyL (PSMA) uptake and the majority of lesions assigned a PSMA score of 1 (≥

blood pool activity and lower than liver uptake). Despite the ongoing debate on the optimal management and timing of surgery in women with advanced HGSOC and conflicting results in various trials (3, 30–32), accurate delineation of disease extent along with several other predictive parameters is crucial for personalizing management, with the goal of offering PCS to women in whom it is feasible to achieve cytoreduction to no gross residual disease (33). Although the results of the current study suggest that ¹⁸F-DCFPyL (PSMA) PET is not a promising modality for imaging of advanced high-grade ovarian cancers, we believe the study protocol developed including detailed comparison of disease mapped on imaging to findings at surgery and surgical histopathology can be utilized in future trials assessing other potential molecular probes targeting receptors or the tumor microenvironment in HGSOC.

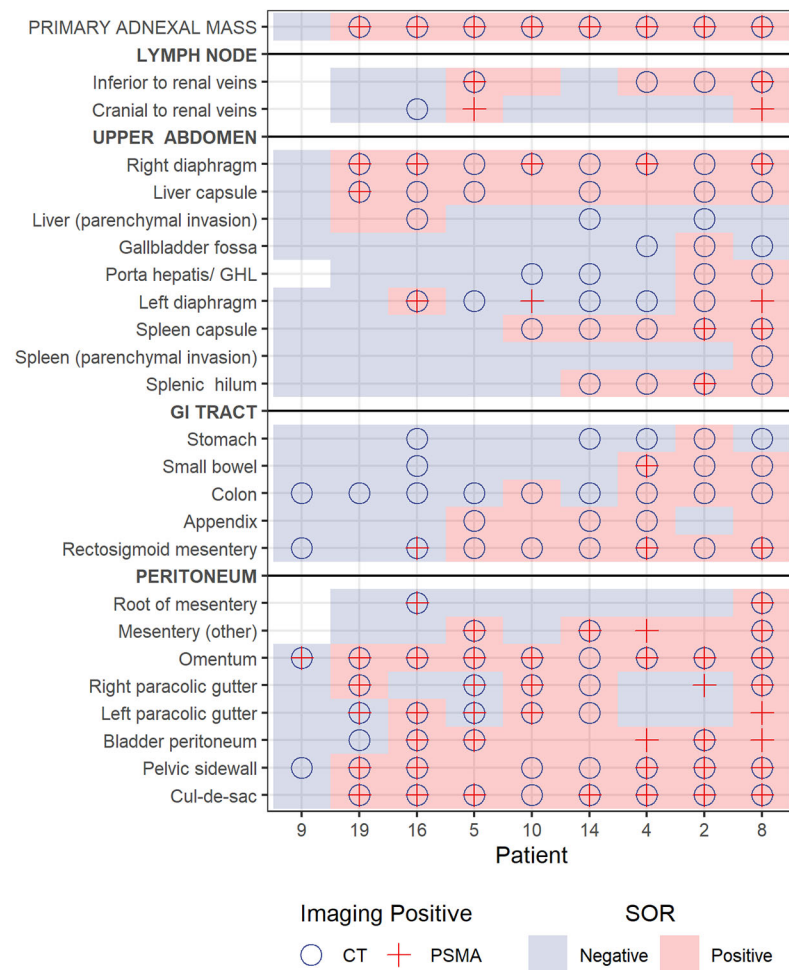


FIGURE 2

Plot showing lesion detection by site for each of the 9 patients who underwent primary cytoreductive surgery. Blue background and pink background denote negative or positive station according to reference standard, respectively. A circle notes positive on CT and cross notes positive on PET.

TABLE 4 Sensitivity and specificity of CT and PET in identifying metastatic sites compared to surgery.

	Sensitivity			Specificity		
	CT (95% CI)	PET (95% CI)	<i>p-value</i>	CT (95% CI)	PET (95% CI)	<i>p-value</i>
LYMPH NODES	0.57 (0.18, 0.90)	0.57 (0.18, 0.90)	1	0.89 (0.52, 1.00)	1.00 (0.66, 1.00)	1
UPPER ABDOMEN	0.88 (0.73, 0.97)	0.32 (0.17, 0.51)	<0.001	0.80 (0.66, 0.91)	0.98 (0.88, 1.00)	0.021
GI TRACT	0.89 (0.67, 0.99)	0.16 (0.03, 0.40)	<0.001	0.54 (0.33, 0.73)	0.96 (0.80, 1.00)	<0.001
PERITONEUM	0.82 (0.68, 0.92)	0.78 (0.63, 0.89)	0.75	0.72 (0.51, 0.88)	0.76 (0.55, 0.91)	1

GI, gastrointestinal; CI, confidence interval.

Bold values are those with statistical significance.

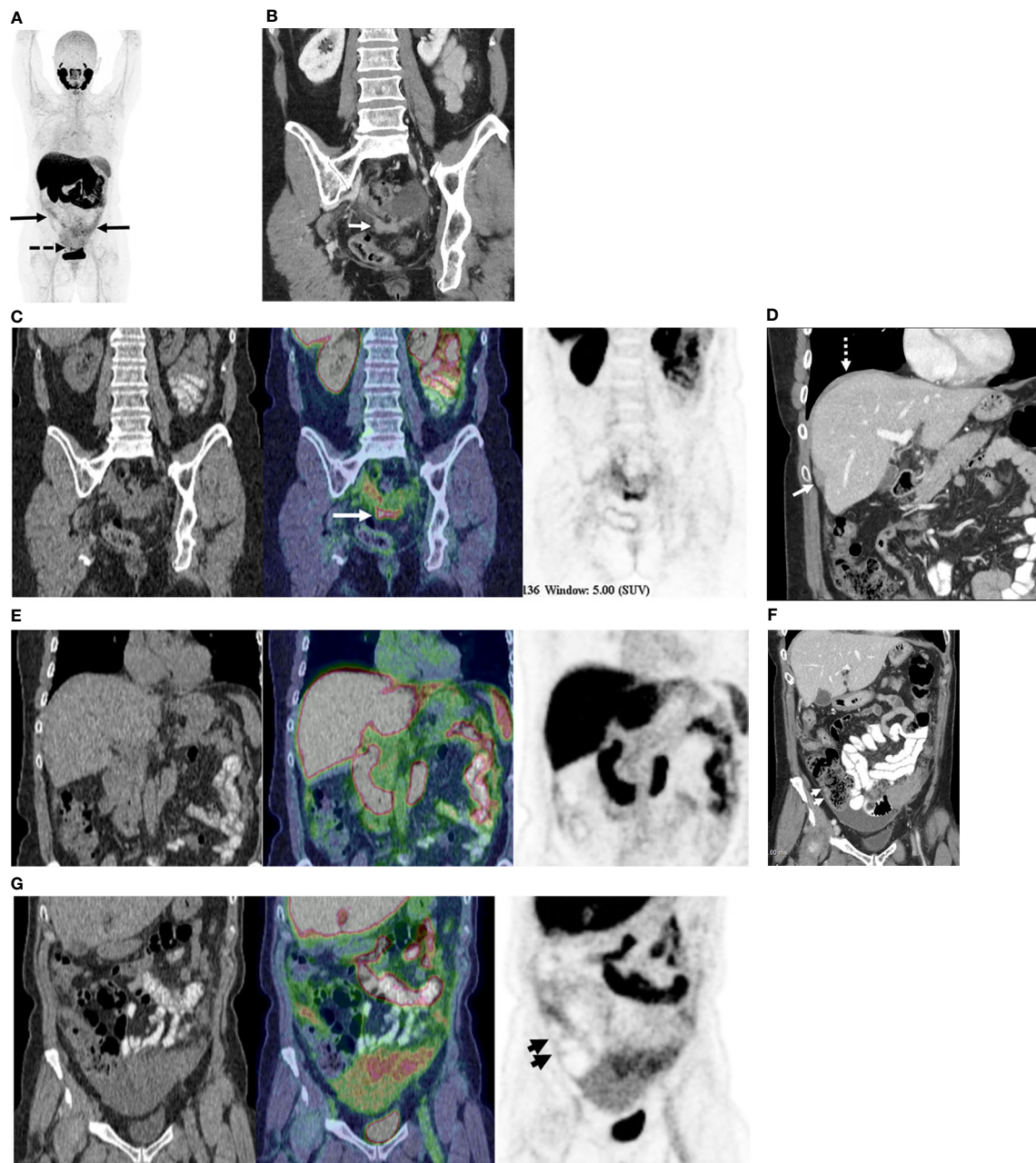


FIGURE 3

68-year old with stage IIIC high-grade serous ovarian cancer. (A) Maximum Intensity Projection image of 18F-DCFPyL shows mild to moderate radiotracer uptake in omental metastases (solid arrows) and moderate radiotracer uptake in pelvic peritoneal metastases (dotted arrows). (B) Coronal contrast enhanced CT image shows peritoneal deposits in the posterior cul-de-sac (arrow). (C) Coronal PET/CT image (CT - left, fused PET/CT image - middle; PET - right) corresponding to B shows moderately radiotracer uptake in same metastatic deposit (concordant CT and PET). Metastatic disease was confirmed at surgery. (D) Coronal contrast enhanced CT image shows metastatic disease on right diaphragm (dotted arrow) and along capsular surface of liver (solid arrow). (E) Coronal PET/CT image (CT - left, fused PET/CT image - middle; PET - right) corresponding to D show no focal radiotracer uptake visible on right diaphragm or liver capsule. Surgical pathology confirmed CT findings of metastatic disease at these sites. (F) Coronal contrast enhanced CT image shows focal thickening along right lateral wall of ascending colon (short arrows), suspected to be represented serosal metastasis. (G) Coronal PET/CT image (CT - left, fused PET/CT image - middle; PET - right) corresponding to F shows no focal radiotracer uptake on the serosal surface of the ascending colon. No serosal disease on surface of the right colon was found at surgery.

Over 80% of patients with advanced HGSOc will experience recurrence within 59 months from initial treatment, with median progression free survival of ~14–15 months (31, 32). One of the proposed mechanisms for the high recurrence rates, is the development of drug resistance to platinum-based chemotherapy, including in patients who were initially responsive to platinum-based chemotherapy protocols. This disease course encourages exploration of new adjuvant therapies to improve disease control and improves outcomes. One example could be utilization of radionuclide therapy in a theranostic approach, where a specific biomarker is employed to image and to deliver targeted radiotherapy selectively to tumor sites. This approach has been effective in a few malignancies including metastatic prostate cancer where the provision of ^{177}Lu -PSMA-617 in men with advanced PSMA-avid metastatic castration-resistant prostate cancer has been shown to improve progression-free survival and overall survival compared to standard care (34). A prerequisite for success of a theranostic approach is high avidity of the biomarker at tumor sites. In addition to demonstrating limited utility of ^{18}F -DCFPyL PET/CT as a diagnostic tool in the staging of women with HGSOc, our findings also suggest that further exploration of PSMA as a target for a theranostic approach is unlikely to be productive in women with advanced ovarian cancers.

In conclusion, although ^{18}F -DCFPyL has higher specificity than CT in detecting advanced HGSOc tumor sites, it detects less disease sites than CT, especially in the upper abdomen and along the gastrointestinal tract, limiting its clinical utility as a diagnostic tool. Further imaging biomarkers with high target affinity are needed to improve disease detection and for a theranostic approach to be considered in women with advanced high-grade epithelial ovarian cancers.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

References

1. Momenimovahed Z, Tiznobaik A, Taheri S, Salehiniya H. Ovarian cancer in the world: Epidemiology and risk factors. *Int J Womens Health* (2019) 11:287–99. doi: 10.2147/IJWH.S197604
2. Bristow RE, Tomacruz RS, Armstrong DK, Trimble EL, Montz FJ survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: A meta-analysis. *J Clin Oncol* (2002) 20:1248–59. doi: 10.1200/JCO.2002.20.5.1248
3. Vergote I, Tropé CG, Amant F, Kristensen GB, Ehlen T, Johnson N, et al. Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *N Engl J Med* (2010) 363:943–53. doi: 10.1056/NEJMoa0908806
4. du Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: A combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials. *Cancer* (2009) 115:1234–44. doi: 10.1002/cncr.24149

Ethics statement

The studies involving human participants were reviewed and approved by UHN Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

UM, LH: conceptualization, methodology, resources and supervision. RK, TC, SJ, LA, DH: investigation and formal analysis. PV-H, MB: writing, review and editing. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.1025475/full#supplementary-material>

5. Rose PG, Nerenstone S, Brady MF, Clarke-Pearson D, Olt G, Rubin SC, et al. Secondary surgical cytoreduction for advanced ovarian carcinoma. *N Engl J Med* (2004) 351:2544–6. doi: 10.1056/NEJMoa041125
6. Dowdy SC, Mullany SA, Brandt KR, Huppert BJ, Cliby WA. The utility of computed tomography scans in predicting suboptimal cytoreductive surgery in women with advanced ovarian cancer. *Cancer* (2004) 101:346–52. doi: 10.1002/cncr.20376
7. Metser U, Jones C, Jacks LM, Bernardini MQ, Ferguson S. Identification and quantification of peritoneal metastases in patients with ovarian cancer with multidetector computed tomography: Correlation with surgery and surgical outcome. *Int J Gynecol Cancer* (2011) 21(8):1391–8. doi: 10.1097/IGC.0b013e31822925c0
8. De Iaco P, Musto A, Orazi L, Zamagni C, Rosati M, Allegri V, et al. FDG-PET/CT in advanced ovarian cancer staging: Value and pitfalls in detecting lesions in different abdominal and pelvic quadrants compared with laparoscopy. *Eur J Radiol* (2011) 80(2):e98–103. doi: 10.1016/j.ejrad.2010.07.013
9. Healy JC. Detection of peritoneal metastases. *Cancer Imaging* (2001) 1(2):4–12. doi: 10.1102/1470-7330.2001.002
10. Han S, Woo S, Suh CH, Lee JJ. Performance of pre-treatment ^{18}F -fluorodeoxyglucose positron emission tomography/computed tomography for detecting metastasis in ovarian cancer: A systematic review and meta-analysis. *J Gynecol Oncol* (2018) 29(6):e98. doi: 10.3802/jgo.2018.29.e98
11. Lopez-Lopez V, Cascales-Campos PA, Gil J, Frutos L, Andrade RJ, Fuster-Quinonero M, et al. Use of (18)F-FDG PET/CT in the preoperative evaluation of patients diagnosed with peritoneal carcinomatosis of ovarian origin, candidates to cytoreduction and hipec. A pending issue. *Eur J Radiol* (2016) 85(10):1824–8. doi: 10.1016/j.ejrad.2016.08.006
12. Metser U, Zukotynski K, Mak V, Langer D, MacCrostie P, Finelli A, et al. Effect of ^{18}F -DCFPyL PET/CT on the management of patients with recurrent prostate cancer: Results of a prospective multicenter registry trial. *Radiology* (2022) 303(2):414–22. doi: 10.1148/radiol.211824
13. Emmett L, van Leeuwen PJ, Nandurkar R, Scheltema MJ, Cusick T, Hruba G, et al. Treatment outcomes from ^{68}Ga -PSMA PET/CT-informed salvage radiation treatment in men with rising PSA after radical prostatectomy: Prognostic value of a negative PSMA PET. *J Nucl Med* (2017) 58(12):1972–6. doi: 10.2967/jnumed.117.196683
14. Roach PJ, Francis R, Emmett L, Hsiao E, Kneebone A, Hruba G, et al. The impact of ^{68}Ga -PSMA PET/CT on management intent in prostate cancer: Results of an Australian prospective multicenter study. *J Nucl Med* (2018) 59(1):82–8. doi: 10.2967/jnumed.117.197160
15. Fendler WP, Calais J, Allen-Auerbach M, Bluemel C, Eberhardt N, Emmett L, et al. ^{68}Ga -PSMA-11 PET/CT interobserver agreement for prostate cancer assessments: An international multicenter prospective study. *J Nucl Med* (2017) 58(10):1617–23. doi: 10.2967/jnumed.117.190827
16. van Leeuwen PJ, Stricker P, Hruba G, Kneebone A, Ting F, Thompson B, et al. (68) Ga-PSMA has a high detection rate of prostate cancer recurrence outside the prostatic fossa in patients being considered for salvage radiation treatment. *BJU Int* (2016) 117(5):732–9. doi: 10.1111/bju.13397
17. Giesel FL, Knorr K, Spohn F, Will L, Maurer T, Flechsig P, et al. Detection efficacy of [^{18}F]PSMA-1007 PET/CT in 251 patients with biochemical recurrence after radical prostatectomy. *J Nucl Med* (2019) 60(3):362–8. doi: 10.2967/jnumed.118.212233
18. Schmidt-Hegemann NS, Stief C, Kim TH, Eze C, Kirste S, Strouthos I, et al. Outcome after PSMA PET/CT based salvage radiotherapy in patients with biochemical recurrence after radical prostatectomy: A bi-institutional retrospective analysis. *J Nucl Med* (2019) 60(2):227–33. doi: 10.2967/jnumed.118.212563
19. Gorin MA, Pienta KJ, Pomper MG, Rowe SP. Prostate cancer local recurrence detected with both ^{18}F -fluciclovine and PSMA-targeted ^{18}F -DCFPyL PET/CT. *Urology* (2017) 107:e9–e10. doi: 10.1016/j.urology.2017.06.015
20. Chen Y, Pullambhatla M, Foss CA, Byun Y, Nimmagadda S, Senthambhichelvan S, et al. 2-(3-{[1-Carboxy-5-[(6-[18F]fluoro-pyridine-3-carbonyl)-amino]-pentyl]-ureido)-pentanedioic acid, [18F]DCFPyL, a PSMA-based PET imaging agent for prostate cancer. *Clin Cancer Res* (2011) 17(24):7645–553. doi: 10.1158/1078-0432.CCR-11-1357
21. Dietlein M, Kobe C, Kuhnert G, Stockter S, Fischer T, Schomäcker K, et al. Comparison of [18F]DCFPyL and [68Ga]Ga-PSMA-HBED-CC for PSMA-PET imaging in patients with relapsed prostate cancer. *Mol Imaging Biol* (2015) 17:575–84. doi: 10.1007/s11307-015-0866-0
22. Basso Dias A, Finelli A, Bauman G, Veit-Haibach P, Berlin A, Ortega C, et al. Impact of ^{18}F -DCFPyL PET on staging and treatment of unfavorable intermediate or high-risk prostate cancer. *Radiology* (2022) 304(3):600–8. doi: 10.1148/radiol.211836
23. Wernicke AG, Kim S, Liu H, Bander NH, Pirog EC. Prostate-specific membrane antigen (PSMA) expression in the neovasculature of gynecologic malignancies: Implications for PSMA-targeted therapy. *Appl Immunohistochem Mol Morphol* (2017) 25(4):271–6. doi: 10.1097/PAI.0000000000000297
24. Nougaret S, Addley HC, Colombo PE, Fujii S, Al Sharif SS, Tirumani SH, et al. Ovarian carcinomatosis: How the radiologist can help plan the surgical approach. *RadioGraphics* (2012) 32(6):1775–800. doi: 10.1148/rg.326125511
25. Eiber M, Herrmann K, Calais J, Hadaschik B, Giesel FL, Hartenbach M, et al. Prostate cancer molecular imaging standardized evaluation (PROMISE): Proposed miTNM classification for the interpretation of PSMA-ligand PET/CT. *J Nucl Med* (2018) 59(3):469–78. doi: 10.2967/jnumed.117.198119
26. Collett D. *Modelling binary data, 2nd ed.* Boca Raton, FL: Chapman & Hall/CRC (1999) p. 45–101.
27. Smith PJ, Hadgu A. Sensitivity and specificity for correlated observations. *Stat Med* (1992) 11(11):1503–9. doi: 10.1002/sim.4780111108
28. Agresti A. *Categorical data analysis, 2nd ed.* New York, NY: John Wiley & Sons (2002) p. 14–27.
29. Aide N, Poulain L, Elie N, Briand M, Giffard F, Blanc-Fournier C, et al. A PSMA-targeted theranostic approach is unlikely to be efficient in serous ovarian cancers. *EJNMMI Res* (2021) 9:11(1):11. doi: 10.1186/s13550-021-00756-z
30. Kehoe S, Hook J, Nankivell M, Jayson GC, Kitchener H, Lopes T, et al. Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): An open-label, randomised, controlled, non-inferiority trial. *Lancet* (2015) 386:249–57. doi: 10.1016/S0140-6736(14)62223-6
31. Onda T, Satoh T, Ogawa G, Saito T, Kasamatsu T, Nakanishi T, et al. Comparison of treatment invasiveness between upfront debulking surgery versus interval debulking surgery following neoadjuvant chemotherapy for stage III/IV ovarian, tubal, and peritoneal cancers in a phase III randomised trial: Japan clinical oncology group study JCOG0602. *Eur J Cancer* (2016) 64:22–31. doi: 10.1016/j.ejca.2016.05.017
32. Fagotti A, Ferrandina MG, Vizzielli G, Pasciuto T, Fanfani F, Gallotta V, et al. Randomized trial of primary debulking surgery versus neoadjuvant chemotherapy for advanced epithelial ovarian cancer (SCORPION-NCT01461850). *Int J Gynecol Cancer* (2020) 30:1657–64. doi: 10.1136/ijgc-2020-001640
33. Tsonis O, Gkrozou F, Vlachos K, Paschopoulos M, Mitsis MC, Zakynthinakis-Kyriakou N, et al. Upfront debulking surgery for high-grade serous ovarian carcinoma: Current evidence. *Ann Transl Med* (2020) 8(24):1707. doi: 10.21037/atm-20-1620
34. Sartor O, de Bono J, Chi KN, Fizazi K, Herrmann K, Rahbar K, et al. Lutetium-177-PSMA-617 for metastatic castration-resistant prostate cancer. *N Engl J Med* (2021) 385(12):1091–103. doi: 10.1056/NEJMoa2107322



OPEN ACCESS

EDITED BY

Aaron Fenster,
Western University, Canada

REVIEWED BY

Madjid Soltani,
University of Waterloo, Canada
Senthilkumar Mohan,
VIT University, India

*CORRESPONDENCE

Jessica Fitzjohn
jessica.fitzjohn@pg.canterbury.ac.nz

SPECIALTY SECTION

This article was submitted to
Breast Cancer,
a section of the journal
Frontiers in Oncology

RECEIVED 15 June 2022

ACCEPTED 12 October 2022

PUBLISHED 02 November 2022

CITATION

Fitzjohn J, Zhou C and Chase JG
(2022) Breast cancer diagnosis using
frequency decomposition of surface
motion of actuated breast tissue.
Front. Oncol. 12:969530.
doi: 10.3389/fonc.2022.969530

COPYRIGHT

© 2022 Fitzjohn, Zhou and Chase. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction
in other forums is permitted, provided
the original author(s) and the
copyright owner(s) are credited and
that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Breast cancer diagnosis using frequency decomposition of surface motion of actuated breast tissue

Jessica Fitzjohn^{1*}, Cong Zhou^{1,2} and J. Geoffrey Chase¹

¹Department of Mechanical Engineering, Centre for Bio-engineering, University of Canterbury, Christchurch, New Zealand, ²School of Civil Aviation, Northwestern Polytechnic University, Xian, China

This paper presents a computationally simple diagnostic algorithm for breast cancer using a non-invasive Digital Image Elasto Tomography (DIET) system. N=14 women (28 breasts, 13 cancerous) underwent a clinical trial using the DIET system following mammography diagnosis. The screening involves steady state sinusoidal vibrations applied to the free hanging breast with cameras used to capture tissue motion. Image reconstruction methods provide surface displacement data for approximately 14,000 reference points on the breast surface. The breast surface was segmented into four radial and four vertical segments. Frequency decomposition of reference point motion in each segment were compared. Segments on the same vertical band were hypothesised to have similar frequency content in healthy breasts, with significant differences indicating a tumor, based on the stiffness dependence of frequency and tumors being 4~10 times stiffer than healthy tissue. Twelve breast configurations were used to test robustness of the method. Optimal breast configuration for the 26 breasts analysed (13 cancerous, 13 healthy) resulted in 85% sensitivity and 77% specificity. Combining two opposite configurations resulted in correct diagnosis of all cancerous breasts with 100% sensitivity and 69% specificity. Bootstrapping was used to fit a smooth receiver operator characteristic (ROC) curve to compare breast configuration performance with optimal area under the curve (AUC) of 0.85. Diagnostic results show diagnostic accuracy is comparable or better than mammography, with the added benefits of DIET screening, including portability, non-invasive screening, and no breast compression, with potential to increase screening participation and equity, improving outcomes for women.

KEYWORDS

breast cancer, screening, diagnostic, digital image elasto tomography, DIET, ROC, bootstrapping

Introduction

Breast cancer is the most frequent cancer and leading cause of cancer deaths in women worldwide (1–4). In 2018 over 620,000 breast cancer related deaths were recorded and both incidence and deaths are expected to rise (5). Early detection is associated with increased survival rates due to cancer being found at an earlier, more curable stage (3, 6). Currently, x-ray mammography is the gold standard for breast cancer screening and is estimated to reduce mortality by up to 25% (7). However, mammography remains controversial due to painful breast compression and invasive radiation exposure (1). Reduced success in women with dense breast tissue (almost 50% of women (8)), arises due to fibroglandular tissue masking the presence of tumors in mammographic images (9–11), further contributing to mammography's radiologist-dependent performance and reducing sensitivity to as low as 27% (12). Because of these drawbacks, mammography is not recommended for younger women (13, 14), where poorer outcomes do not outweigh the risks, creating inequity of breast care for younger women.

Although mammography is accepted as a large scale screening tool, a number of issues have led mammography sensitivity to be significantly overstated (8). Hollingsworth suggests many mammography studies use a cohort inclusive of palpable tumors, which are larger and therefore expectedly easier to diagnose (15, 16). Similarly, sensitivity calculations based on prevalence screens only (the initial screen) result in a disproportionate number of larger tumors and consequently an inflated value for sensitivity (8, 17).

Further, many studies use interval cancers, cancers which occur following a negative mammogram, but before the next round of screening (18), as a measure of false negatives. This methodology is also flawed, as while some cancers may have begun and developed between screens, slower growing tumors may not present in the screening interval, and consequently would be diagnosed as true positives in the following screen, despite being missed previously. This issue results in sensitivity dependent on screening interval and higher sensitivity than studies comparing supplemental imaging modalities. This method of assessment will also impact specificity values, as false negatives in past mammograms would be included as a true negative reading. In other words, studies which exclude the use of a complementary modality assume mammography diagnosis is true until proven false with only future mammograms to compare to. Even then, cancer found in subsequent mammograms are considered true positives and assumed to have began and developed between screens. This process is clearly methodologically over-simplistic and unsound, and accuracy estimates of sensitivity or specificity using these methods should be discounted appropriately.

Digital Image Elasto Tomography (DIET) (19–23) is an alternative breast cancer screening technology. The system is portable with non-invasive testing, and thus able to increase equity for young women and those living rurally. DIET involves

a patient lying face down while a mechanical actuator induces steady state sinusoidal vibration in the free-hanging breast. Five digital cameras surrounding the breast capture images of the surface vibration at different stages using synchronized strobe lights. These images are converted into displacement data for over 14,000 reference points using surface volume and optical flow techniques by Tiro Medical (Christchurch, New Zealand). Cancer diagnosis using this surface motion data is based on shear wave transmission differences reflected in the surface motion highlighting the differences in underlying tissue stiffness and damping. A significant contrast in elastic properties (400–1000%) between healthy and cancerous tissue provides a potentially highly distinguishable diagnostic, which is much greater than the 5–10% contrast in radio density used by mammography (24, 25). Thus, the DIET system detects and localises a tumor location based on identification of the higher underlying tissue stiffness, which is very different to typical internal imaging modalities.

Analysis of this data has showed potential for diagnostic success including Zhou et al's study on hysteresis loop analysis (HLA) (26, 27) and Kashif et al's study on modal analysis (22). Most diagnostic methods developed were exclusively tested on silicone phantom data, used in early technology development (28). Despite sophisticated silicone phantom design (23, 29), silicone phantom breasts with stiffer inclusions cannot truly imitate the inhomogenous fibrous structure of breast tissue nor the complex interactions between tumors and healthy tissue. Zhou et al's HLA study underwent limited testing with clinical data (3 subjects) (27), but sensitivity was not high and dependent on selection of actuator input frequency for different subjects, showing inconsistent performance.

This author presented one diagnostic method using DIET displacement data, which was validated on 26 breasts (30). This method involved fitting a viscous damping model (VDM) to viscous damping distribution in different breast segments and comparing model coefficients. One model coefficient, related to stiffness showed diagnostic insight with optimal sensitivity and specificity of 77%, using clinical data.

This paper describes a fully automated, computationally efficient diagnostic algorithm, which uses displacement data from a novel 3D surface motion reconstruction DIET technology. The proposed method is based on the hypothesis stiff tumors will affect response frequencies in the breast compared to other regions containing healthy tissue, thus enabling transformation of dynamic response into a novel diagnostic metric to identify regions of higher stiffness for cancer diagnosis. This diagnostic uses a combination of frequency components analysis, surface segmentation and bootstrapping techniques, while previous diagnostic methods based on DIETs technology mainly identify damping, stiffness and modal distribution (22, 27, 30). In addition, it provides an unbiased diagnostic criteria to ensure each breast to be diagnosed independently, regardless of varying breast

properties across the population, which is critical for improving the equity of screening with this technology. Overall, this work offers a novel approach to implement automated unbiased tumor detection in this DIET screening technology.

Method

Clinical data

Fourteen women (P1-P14) were recruited to undergo testing using a prototype DIET system, as part of a clinical trial run at Canterbury Breastcare (Christchurch, New Zealand). Thirteen women had a tumor in one breast and one woman had two healthy breasts resulting in a total of 13 cancerous, 15 healthy breasts. Patient P6 also has an additional non-malignant cyst in their right (healthy) breast. Each woman underwent mammography screening prior to testing using the DIET prototype and diagnostic capabilities of the algorithm presented in this paper aim to correctly distinguish between healthy and cancerous breasts in this cohort and match the diagnostic given by mammography. Ethics approval for the experimental tests, data collection, and analysis of this data was granted by the NZ National Health and Disability Ethics Committee, South Island Regional Committee.

Table 1 shows the patient age, tumor size and location from mammography reports for each patient with a cancerous breast, as well as the approximate breast volume of the cancerous breast calculated using the DIET measured displacement data. Tumor sizes ranged from 7 to 48mm and displacement data was available for a range of testing frequencies (20–50 Hz). While clinical data is limited to 28 breasts, the variation in breast properties and tumor sizes is large, providing a varied cohort. Ensuring diagnostic performance is robust to these variations in tumor size and breast properties is a particular focus in this paper. The accuracy of the size

and location vary and can be difficult to determine from mammography images. Thus, location and size data were treated as approximate, and algorithm success focused on correct diagnosis, rather than precise tumor localisation.

It is important to note patient P14, with two healthy breasts, originally had their right breast, P14R, diagnosed as cancerous, which was later discovered to be healthy tissue. This result shows a false positive in mammography and correct diagnosis of this breast using DIET would further demonstrate its diagnostic potential. Occasionally, difficulties in optical flow or image reconstruction resulted in a lack of data for some subjects or a limited number of available input frequencies. The algorithm presented in this paper uses low actuator input frequencies, resulting in the exclusion of two healthy breasts in P4 and P13. The result is 13 cancerous and 13 healthy breasts used in this analysis and the diagnostic algorithm presented in this paper aims to correctly diagnose these subjects.

Diagnostic criteria

As mentioned, sensitivity and specificity of mammography has been overstated in many studies. To assess approximate true sensitivity and specificity, studies were considered if diagnostic results of mammography were validated using another modality, such as ultrasound or MRI. Values for both dense and non-dense breasts were used when studies distinguished between the two, based on approximately 50% of women having dense breasts (8). Average sensitivity of the ten studies assessed was 60% (40%–78% range) and average specificity was 80% (46%–99% range) (31–40). These accuracy values are more suitable for comparing mammography to other breast screening modalities.

The area under a receiver operator characteristic (ROC) curve (AUC) between 0 and 1 is commonly used to compare diagnostic methods, as a higher AUC indicates a better optimal

TABLE 1 Patient age, breast volume, tumor size and locations cancerous breasts where 12 o'clock is the top of the breast.

Subject Number	Age	Cancerous Breast	Tumor Location (around breast)	Tumor diameter (mm)	Tumor depth (mm)	Cancerous breast Volume (cm^3)
P1	50	Left	10 o'clock	18	35.1	691.7
P2	58	Left	2.30 o'clock	15	–	307.9
P3	58	Right	10 o'clock	14	51.7	265.8
P4	37	Left	3-5 o'clock	48	–	740.8
P5	38	Left	12.30 o'clock	14	26.3	239.5
P6	45	Left	6 o'clock	12	13.3	708.6
P7	55	Left	2 o'clock	23	78.8	673.7
P8	51	Right	10.30 o'clock	37	83.9	1022.0
P9	55	Left	12 o'clock	16	53.1	474.4
P10	50	Left	9-12 o'clock	7	–	1057.0
P11	51	Right	10 o'clock	7	–	342.9
P12	55	Left	11 o'clock	10	–	444.2
P13	47	Right	9-3 o'clock	18	–	593.8

sensitivity and specificity. AUC greater than 0.7, 0.8 and 0.9 are considered acceptable, excellent and outstanding, respectively (41). The average ROC curve AUC value for mammography across eight studies was 0.73 (0.54–0.84 range), which will also be used to compare diagnostic success of the method presented in this paper (12, 31, 40, 42–46).

Two overall accuracy criteria are suggested to show diagnostic potential for the algorithm presented:

1. *Diagnostic sensitivity and specificity similar to mammography (60% sensitivity, 80% specificity)* Achieving comparable diagnostic accuracy to mammography would allow the DIET technology to realise its many benefits including comfort, portability and safety for all women without compromising on diagnostic success.
2. *A highly sensitive diagnostic algorithm is achieved (80% sensitivity, 65% specificity)* An algorithm capable of providing sensitivity higher than mammography will be considered a success even if specificity is slightly lower. This criteria is due to the ease of DIET testing and its many other benefits making it an attractive solution for breast screening. There is potential for added clinical breast exam (CBE) or other breast screening technologies with higher specificity to optimise diagnosis and reduce false positive biopsies, following a highly sensitive diagnosis using DIET.

To maximise the benefit of the DIET technology, this diagnostic algorithm should meet the criteria in Table 2.

Stiffness dependent vibration

The stiffness dependence of vibration frequency is well documented (47). Stiffer materials vibrate at higher frequencies, based on:

$$\omega = \sqrt{\frac{k}{m}} \quad (1)$$

where ω is frequency, k is stiffness and m is mass. Thus, the presence of tumors, known to be 4–10 times stiffer than healthy tissue

[24,48,49,50], may result in a visible increase in response frequency, given mass at a local point is similar, yielding:

$$\omega_{\text{tumor}} = \sqrt{\frac{k_{\text{tumor}}}{m}} = \sqrt{\frac{(4 \sim 10) k_{\text{healthy}}}{m}} = 2 \sim 3.16 \omega_{\text{healthy}} \quad (2)$$

This equation further suggests higher frequency response in cancerous breasts, or breast segments, of a magnitude 2–3 times the response of healthy tissue. Development of a diagnostic algorithm using DIET concluded variation of breast properties, including stiffness and viscous damping, across the population can exceed variations between cancerous and healthy tissue in an individual (30). Therefore, it is important to understand single diagnostic thresholds for response frequencies may not be suitable for diagnosis across a varied cohort and may cause women with naturally stiff breasts to experience a disproportionate number of false positive diagnoses. The author's prior work presented a breast segmentation methodology to analyse tissue properties in different regions of a breast (30). Healthy breasts were hypothesised to have similar breast properties in different segments and larger discrepancies were indicative of a tumor. Applying this methodology, it is hypothesised frequency response of different areas in a healthy breast will be similar; in contrast, tumor presence is expected to affect response frequency.

Frequency component of each reference point vibration response

Displacement data for over 14,000 reference points on each breast surface was provided by Tiro Medical (Christchurch, New Zealand) following clinical testing using the DIET system at each input frequency for the 26 breasts from 14 patients in Section 2.1. The Fourier transform of each reference point signal was implemented in *Matlab* (48) to obtain frequency components of each reference point vibration. Magnitudes of frequency components were ordered, with the dominant frequency expected to be equivalent to the actuator input frequency at the induced steady state response. Based on the knowledge of high mechanical stiffness resulting in higher frequency of response, and cancerous tissue resulting in 400–1000% higher stiffness than healthy tissue, the second dominant frequency has the potential to provide diagnostic information. This latter frequency is hypothesised to be higher, but, more importantly, different in regions of the breast containing a tumor.

The second dominant frequency and its signal magnitude were obtained for each reference point. Reference points with second frequency magnitudes less than 15% of the dominant frequency magnitude were discarded to avoid using reference points where frequency composition was highly varied and the second frequency not considered particularly dominant.

Frequency composition is considered irregular when the dominant frequency is not equivalent to the input frequency. One reason for this irregularity is noise near the actuator or chest

TABLE 2 Diagnostic criteria to assess the success of diagnostic algorithms using DIET.

1. Unbiased diagnostic, unaltered by known tumor identification or symptoms
2. Full automation with no human interpretation of results required
3. Ability to diagnose tumors down to 7mm, the smallest tumors in this clinical data
4. Robust to varying breast sizes and densities
5. Have AUC greater than mammography (>0.73)
6. Meet one of the following accuracy criteria:
 - a. Sensitivity (60%) and specificity (80%) similar to mammography
 - b. Sensitivity (80%) and specificity (65%) highly sensitive screening tool

wall, the latter potentially due to issues of wave reflection. This issue was removed by removing the top and bottom 5% of points, as these areas tend to result in the most noise in both cancerous and healthy breasts (30). While necessary to remove these irregularities, acknowledging they are likely to be a source of increased false positives, it is important to limit the number of points discarded, to avoid tumors close to the nipple or chest wall being missed. Tumors developing near the chest wall are also an existing challenge in mammography due to difficulties in obtaining sufficient breast compression close to the chest wall to allow x-ray penetration (49).

Furthermore, irregular frequency composition could be the result of irregular vibrations for points centred near breast concavities, which are unable to be consistently removed in the current DIET image processing. Alternatively, and most importantly, it could be a result of highly variable breast tissue properties, such as stiff cancerous lesions, having a significantly large effect on frequency composition, showing high diagnostic potential. For the latter reason, reference points with dominant frequency not equal to input frequency were not excluded, and, in these cases, the dominant frequency, rather than the second dominant frequency is used as the frequency of interest. The result is a *frequency of interest* for each reference point. Figure 1 shows a flowchart of this selection process.

Because reference points with dominant frequency not equal to input frequency could be the result of noise and breast concavities, it

was necessary to manage the number used in analysis. At higher input frequencies, average dominant frequencies in segments of the breast differed to the input frequency (>2 Hz difference), suggesting high numbers of irregular points. Thus, displacement data from lower actuation input frequencies was used (20 to 23 Hz) in this analysis, as they resulted in less than 10% of segments exhibiting this irregular trend, compared with 11–39% of segments in frequencies above 23 Hz.

Breast segmentation and unbiased diagnostic criteria

Three-dimensional (3D) colour plots of the breast showing frequencies of interest are presented for three subjects at an input frequency of 23 Hz to show more regions of high response frequencies in cancerous breasts, demonstrating how comparison of these values could provide useful diagnostic insight.

To implement an unbiased diagnostic algorithm utilising this *frequency of interest*, the breast was segmented into four radial segments and four vertical (z) segments, a total of 16 segments (Figure 2). Frequencies of interest were averaged for all reference points in each segment and mean values averaged across available frequencies from 20–23 Hz. Each z-band was analysed separately and one of the four radial segments

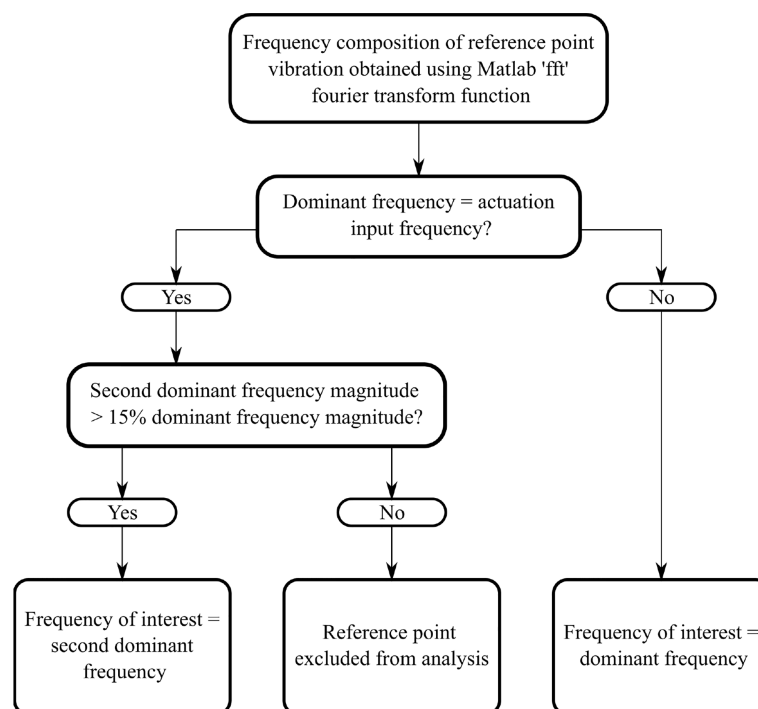


FIGURE 1
Flowchart showing method to obtain frequency of interest for each reference point.

identified as the control. The mean frequency of interest for this control segment was plotted against the mean frequencies of interest for the three other segments in the same z-band. This process was repeated for each z-band, resulting in a total of 12 data points per breast. Occasionally, all reference points in a segment were excluded, based on insufficient magnitude of the second dominant frequency (Figure 1), resulting in less data points per subject.

Different percentage tolerances were used to analyse the degree of similarity between these averaged response frequencies in each separate z-band. Healthy breasts are hypothesised to have similar response frequencies remaining within the tolerance. In contrast, the presence of any one segment outside tolerance levels suggests a cancerous diagnosis. This diagnostic criteria is shown in Figure 3. To test robustness to tumor location 12 breast configurations were trialled with different control segments. Figure 2 shows these 12 control segment configurations for the left breast. The right breast segments were the inverse of these configurations to compare outer and inner breast properties, consistently.

This segmentation process provides a method for each breast to be diagnosed independently, removing issues of highly variable breast properties across the population, within breasts of the same women and due to breast changes over time (50–55). Segmentation both radially and vertically is expected to improve diagnostic outcomes for smaller tumors, whose properties may be more easily distinguishable in a smaller segment.

ROC curve and bootstrapping

ROC curves presenting different percentage tolerances were used to test the sensitivity of this method to predefined diagnostic tolerance levels and assess whether the criteria outlined in Table 2 could be met. ROC curves were used to find both optimal breast configuration and tolerances, which result in criteria being met for sensitivity and specificity of this

method. The discrete ROC curve for two opposite breast configurations are shown, as well as bootstrapped curves for all configurations.

Bootstrapping is used to up-sample data and involved selecting 50 breasts with replacement from the 26 breast cohort. This selection was repeated 200 times and the varying sensitivity and specificity recorded for a number of percentage tolerance thresholds for each trial. A line of best fit was fit to the compounded points of every trial using $y=1-e^{-ax}$ using total least squares. This equation form is able to capture the linear (50:50 chance) line and, with a very large exponent, the perfect square ROC curve, as well as all likely shapes in between. It thus provides a good approximation of the diagnostic performance of each configuration in an ideally larger cohort of data and can be used to assess the performance of this algorithm against the criteria in Table 2. Optimal accuracy, as well as 80% sensitivity and specificity points used to assess accuracy criteria, are marked on the ROC curves. ROC curve AUC was also assessed to ensure it meets the criteria ($AUC > 0.73$). Figure 4 shows a flowchart of the combined methods used to generate the results presented in this paper.

The results in the following section show:

1. 3D plots showing frequency of interest for three breasts at actuator input frequency of 23 Hz, showing proof of diagnostic theory with larger discrepancies and regions of high response frequency in cancerous breasts
2. Unbiased, clinically feasible diagnosis with percentage tolerance used to determine the degree to which more dominant frequencies are different amongst segments in the same breast for both healthy and cancerous breasts for breast configurations 1 and 6
3. Identification of breast and tumor characteristics of false negative subjects
4. Identification of patient age and breast size for false negative and false positive subjects

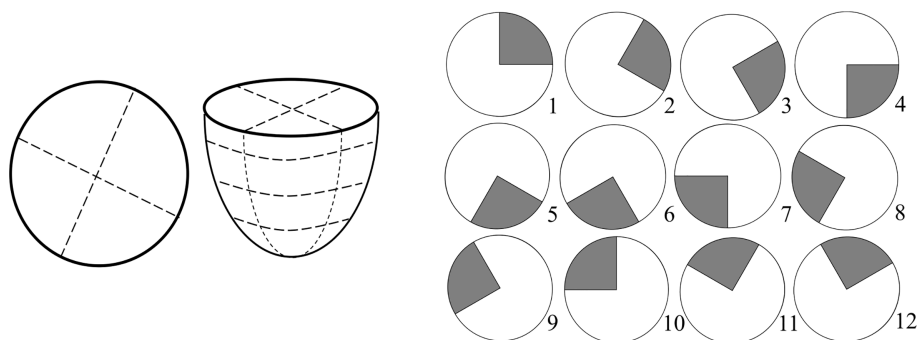


FIGURE 2

Schematic of breast segmentation including vertical (z) segmentation into four bands for a total of 16 segments (left) and diagram of 12 control segment configurations for left breast used to test robustness (right).

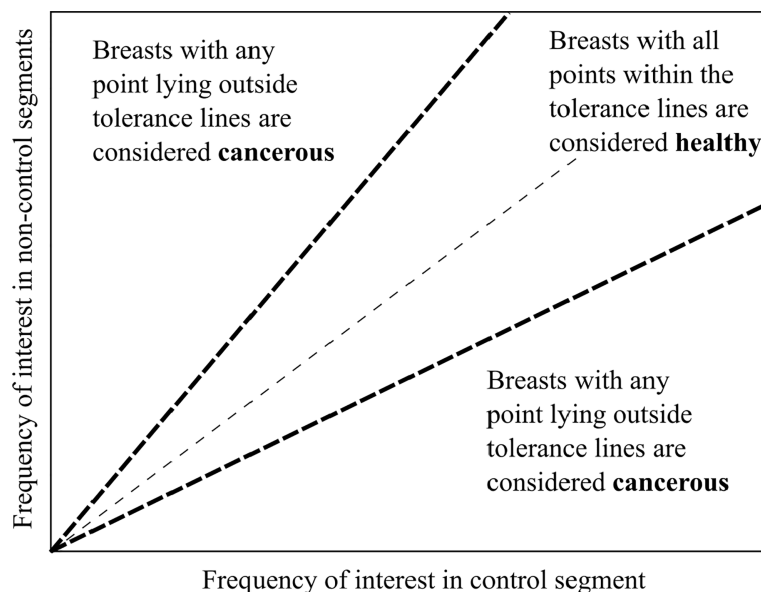


FIGURE 3
Tumor acceptance criteria showing simple diagnostic method.

5. A discrete ROC curve with sensitivity and specificity shown for each percentage tolerance for breast configurations 1 and 6 across all subjects
6. Smooth bootstrapped ROC curves for all breast configurations showing optimal sensitivity and specificity and points at 80% specificity and 80% sensitivity (points of interest for assessing diagnostic success against specified criteria)
7. Table outlining optimal sensitivity, specificity and assessment of each configuration against diagnostic criteria
8. Table assessing this diagnostic method against all criteria outlined in Table 2

Results

Frequency response distribution

Figure 5 shows 3D plots of the distribution of frequencies of interest identified in Section 2.4 for three subjects, showing high frequencies, twice the input frequency and above. The plots clearly show a distinguishable difference in terms of frequency components of healthy and cancerous breasts on the left and right, respectively.

It is important to note, while these images show a significant contrast in frequency, and could potentially provide successful diagnosis based on image observation, direct image observation would result in human assessment of results and a lack of automation. Equally, such observation could be used to

reinforce or check any automated diagnostic. Thus, these images show a proof-of-concept justification for using frequency composition to infer diagnosis, but require further development of unbiased, algorithm automation shown in consequent sections of this paper, to be clinically feasible.

Unbiased, clinically feasible diagnosis

Figure 6 shows the diagnostic result of applying optimal percentage tolerance 34% using optimal configuration 6 and optimal percentage tolerance 33% for breast configuration 1 for both cancerous and healthy breasts. This figure shows false negatives may be dependent on configuration orientation, likely due to varying tumor locations and tumors effecting segments on either side. These two configurations, positioned in the upper outer and lower inner portions of the breast, respectively, demonstrate using the result of two separate configurations ensures all cancer is diagnosed.

Figure 7 shows breast and tumor characteristics for the false negative subjects identified in configurations 1 and 6. The figures show tumor size compared to breast volume and depth (normalised by volume), respectively for each configuration. Figure 8 shows patient age and breast size for false negatives and false positives, respectively.

Figure 9 shows discrete ROC curves for diagnostic performance at different percentage tolerances for configurations 6 and 1. Figures 10, 11 show the bootstrapped ROC curves for all breast configurations with optimal points shown and Table 3 shows the resulting AUC and assessment against diagnostic criteria.

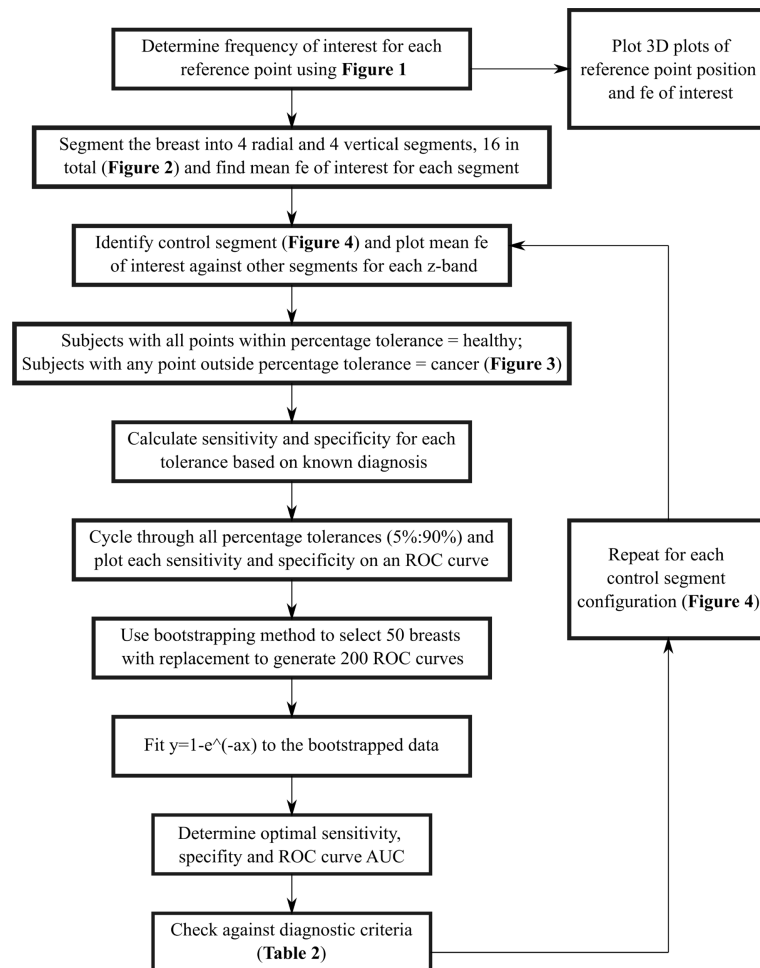


FIGURE 4
Flowchart showing full method to obtain results.

Assessment against diagnostic criteria

The performance of this algorithm was assessed against the diagnostic criteria in Table 2, as shown in Table 4. All criteria were met and each 7 mm tumor was correctly diagnosed in one of the breast configurations analysed. Diagnosis using this algorithm is unbiased and completely automated and diagnostic accuracy exceeds mammography, showing significant diagnostic efficacy across this varied cohort.

Discussion

Proof of concept

Figure 5 shows a visual representation of how the frequency of interest is higher and more varied in cancerous breasts.

Healthy breasts generally show more areas of purple, suggesting lower frequencies and less sections of high frequency content. In contrast, cancerous breasts are generally seen to exhibit larger areas of high frequency response, which is expected based on stiffer materials vibrating at higher frequencies and cancerous tissue having stiffness 4~10 times greater than healthy tissue (24, 56–58). The examples in Figure 5 were typical for most subjects at low frequencies (<26 Hz). Despite this significant contrast, observation of stiffness plots alone cannot quantify diagnosis, as it would fail the criteria of being automated and unbiased. These results provide proof-of-concept for the governing theory of this algorithm, but require development to prevent reliance on human interpretation, which would increase bias and error.

Furthermore, the notable differences seen here were not consistent throughout all frequencies. In general, higher frequencies (>26 Hz) resulted in much noisier and varied

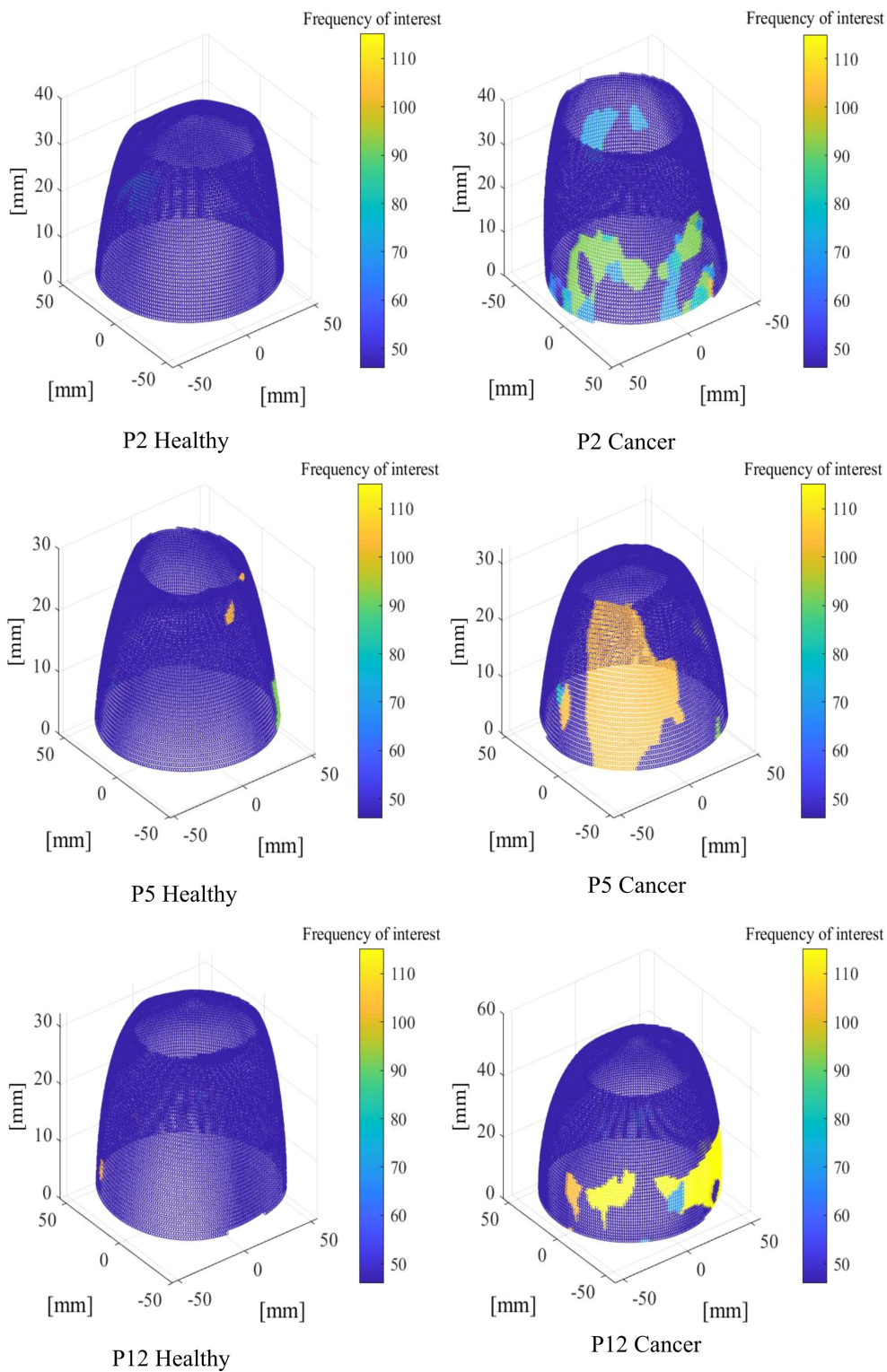


FIGURE 5
Three-dimensional (3D) plots showing areas of high frequency of interest for healthy (left) and cancerous (right) breasts at input frequency, $f_e = 23$ Hz.

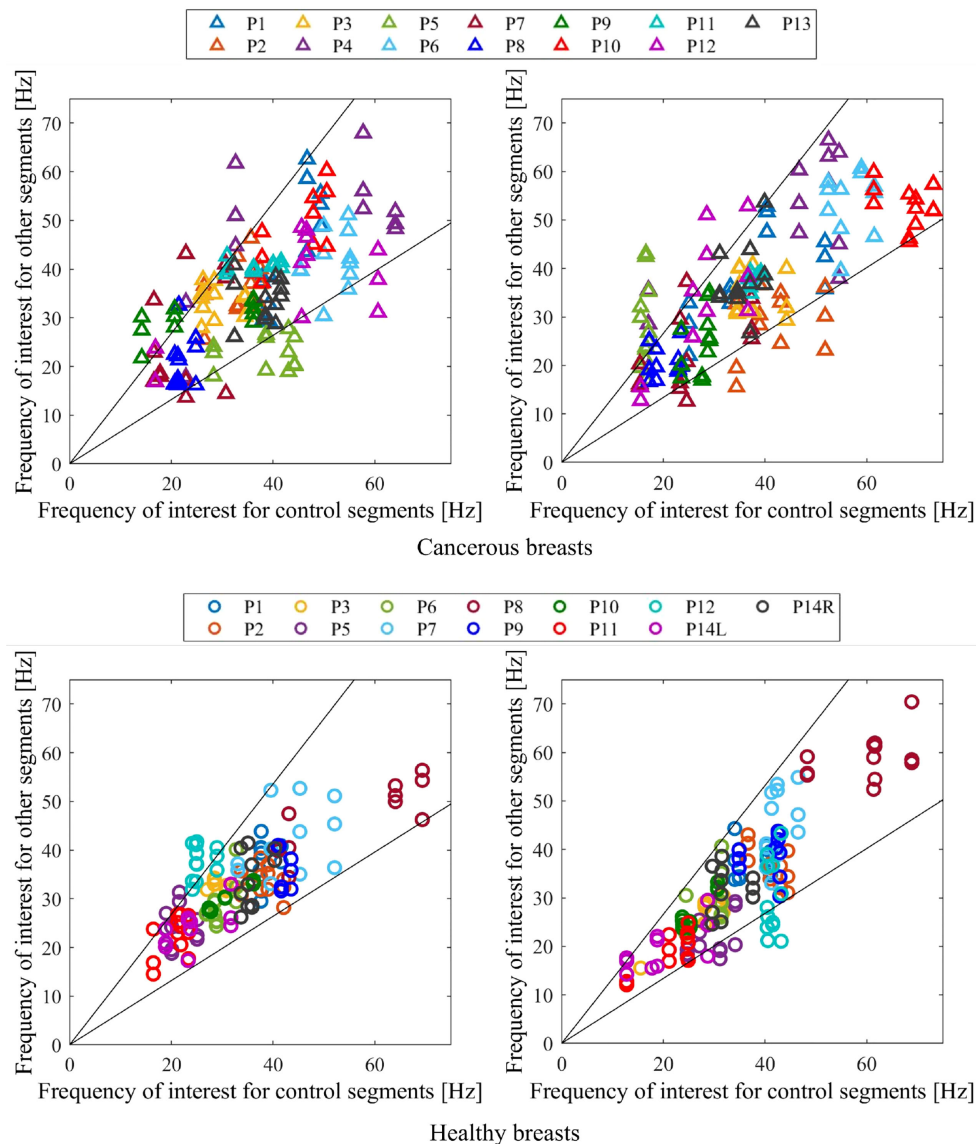


FIGURE 6

Diagnostic criteria for cancerous (top) and healthy (bottom) breasts for configuration 6 with 34% tolerance applied (left) and configuration 1 with 33% tolerance applied (right). Any one point lying outside the percentage tolerance shown results in a cancerous diagnostic.

frequency content across all breasts, likely due to a higher incidence of wave reflection causing an increase in vibration in certain areas. As mentioned, higher frequencies tended to result in more atypical vibration with dominant frequency not equal to the input frequency in many cases. Thus, lower frequencies (20–23 Hz) were used for this frequency response analysis. Averaging the frequency of interest across the available frequencies from 20–23 Hz is unbiased and generalisable, although it should be noted, with more subject information, such as breast density, able to be found prior to screening, more optimal breast-specific testing frequencies may possibly be obtained.

Unbiased diagnostic technique

Figure 6 shows how this frequency of interest can be used as an indicator for cancer in an unbiased and clinically feasible way, by using a segmentation methodology and comparing frequencies of interest in different segments of a breast. Figure 6 shows the diagnostic result for two breast configurations (1 and 6 in Figure 2) situated on opposite sides of the breast at their respective optimal tolerances, 33% and 34%. This figure shows considerable variation in frequency composition of segments in cancerous breasts compared to

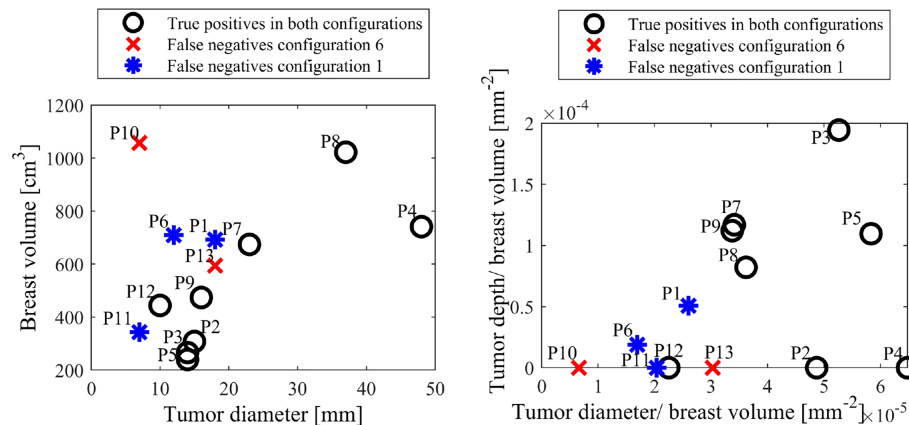


FIGURE 7

False negative tumor and breast sizes (left) and tumor depth and diameter normalised by breast volume (right) for configurations 1 and 6.

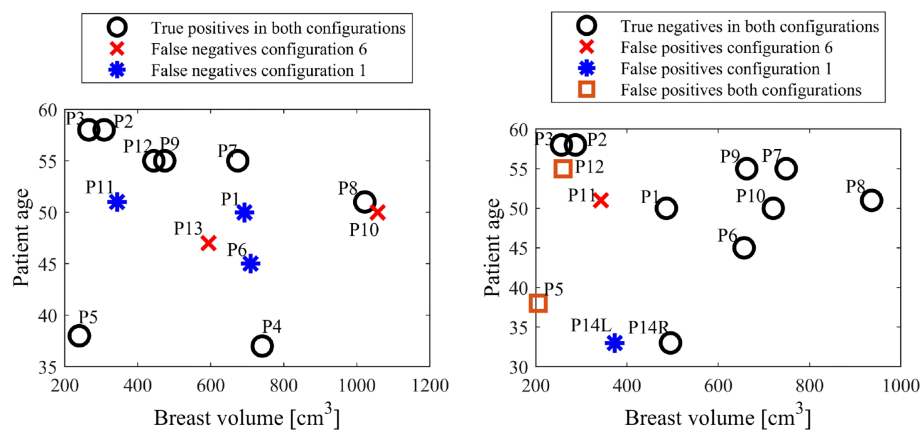


FIGURE 8

False negative (left) and false positive (right) ages and breast volumes for configurations 1 and 6.

healthy breasts. This clear, observable difference supports the use of this diagnostic segmentation methodology, demonstrating how tissue properties in a healthy breast tend to be more similar, as expected. The large variation in average frequencies of interest for both cancerous and healthy breasts in Figure 6 further demonstrates breast properties are unique, and vary even between breasts of the same women, showing set diagnostic thresholds or comparison between breasts is likely to result in inaccurate diagnosis and poor overall performance.

Of particular interest is the varying performance of configurations 1 and 6 in diagnosing specific subjects. Configuration 6 was the optimal configuration and resulted in two false negatives, P10 and P13 (85% sensitivity), and three

false positives, P5, P11 and P12 (77% specificity), already meeting Criteria #2 for a highly sensitive diagnostic algorithm in Table 2. In contrast, configuration 1 resulted in three false negatives, P1, P6 and P11 (77% sensitivity) and three positives, P5, P12 and P14L (77% specificity). Thus, while false positives P5 and P12 were diagnosed incorrectly in both configurations, all false negative diagnoses were diagnosed correctly in one of the two configurations.

Differing diagnostic success for different subjects in each configuration shows tumor location potentially affects the efficacy of diagnosis in certain configurations. Fitzjohn et al. suggests tumor presence can often affect the properties of segments either side (30), and, as such, using these segments

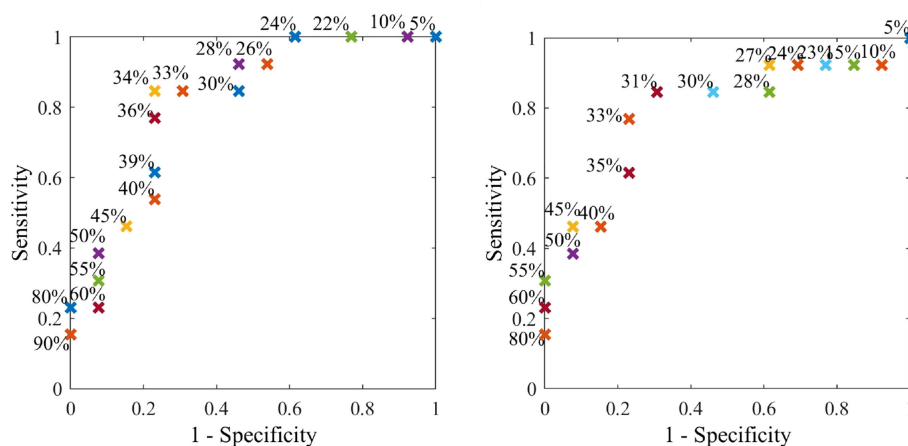


FIGURE 9

Discrete ROC curve showing diagnostic method applied at different percentage tolerances for breast configuration 6.

as the control may result in less distinguishable properties compared to a segment far from the tumor, where a greater difference will result in a more prominent cancer diagnostic.

Configurations 1 and 6 are situated on opposite sides of the breast and, when combined, result in all cancer being diagnosed in at least one segment. Therefore, there is potential for two opposite segments to be used to ensure diagnosis of all cancers. If diagnosis in either configuration was to result in positive diagnosis, the diagnostic result would be zero false negatives (100% sensitivity) and four false positives (69% specificity), meeting Criteria #2 in Table 2, with perfect sensitivity and still acceptable specificity (>65%). This outcome shows a significant diagnostic using a computationally efficient algorithm. Further metrics could be designed to potentially combine with other DIET diagnostic methods and reduce false positive results.

Combining results may improve sensitivity but increase false positives and unnecessary biopsies, which already impact almost 20% of women (59). Clinical breast examination or other breast screening tools may also be utilised to ensure unnecessary breast biopsies are reduced. In particular, positive DIET results could be immediately followed up with skilled manual palpation or ultrasound to reduce this risk and reduce the time taken to women receiving diagnostic outcomes and consequent treatment.

Figure 7 shows different patient, breast and tumor information for the false negatives in both configurations 1 and 6. It shows all false negatives are less than 20 mm, which is associated with lower stage cancer (60) and expectedly considered more difficult to diagnose. Two of the false negatives are the two smallest tumors in this cohort, at 7mm. Most importantly, all cancers are detected in one of the configurations, showing the capability of detecting both 7mm tumors, depending on breast configuration.

Figure 7 also shows all five false negatives across each configuration are five of the six smallest tumor to volume ratios, expected to be more difficult to diagnose in methods comparing average breast segment properties. Additionally, P10, P11 and P13 have unknown tumor depth, which, if deep, could also cause diagnostic issues (30). Figure 8 shows patient age and breast volume for false negative and false positive cases. False negative results occur at a range of breast sizes and average ages for this cohort. More importantly, the true positives are patients with varied ages and breast sizes, suggesting there is no diagnostic limitation of age or volume related breast properties for this algorithm, showing an equitable diagnostic result.

Figure 8 also shows cancer found in two of the youngest women diagnosed correctly, which is a significant result, given mammography often performs worse in young women, who tend to have higher breast density consisting of more glandular tissue, which can mask the presence of a tumor (52, 61). False positives in Figure 8 tend to occur in smaller breasts, perhaps where differing breast structure has a more magnified effect due to smaller segments overall. Thus, adjusting the number of z-bands or segments used based on breast size could potentially reduce false positives.

False positive results could be the result of some complex internal tissue differences around the breast causing distinguishing properties when using this breast segmentation methodology. However, it is important to note a missed diagnosis in mammography should not be ruled out. Specifically, the false positive patients all have smaller breasts, also associated with potentially increased breast density, and thus, worse outcomes in mammography (62). Furthermore, Patients P5 and P14L are the two youngest patients in this cohort. Generally, breast density decreases with increasing age

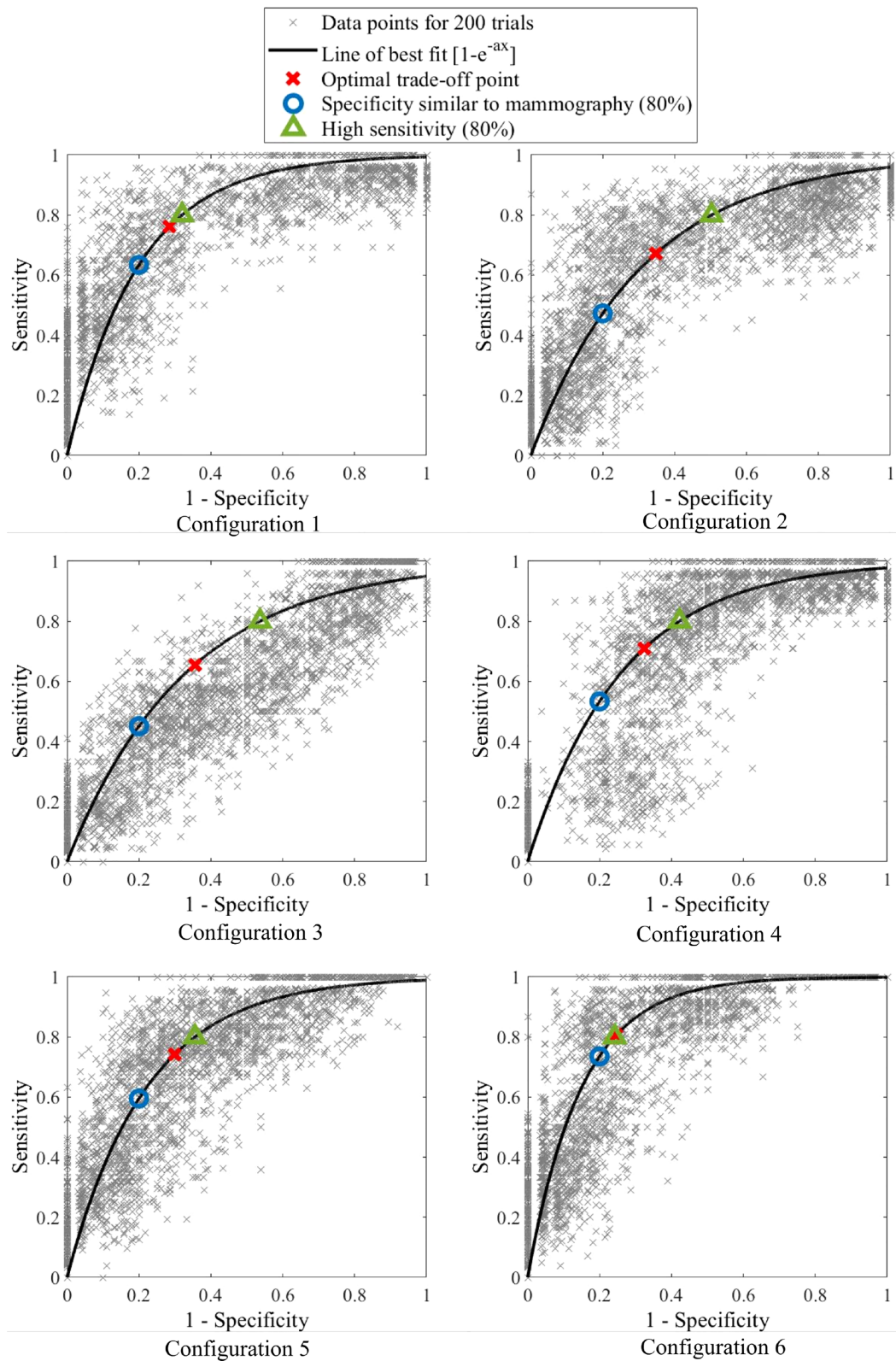


FIGURE 10
Bootstrapped ROC curves for breast configurations 1-6.

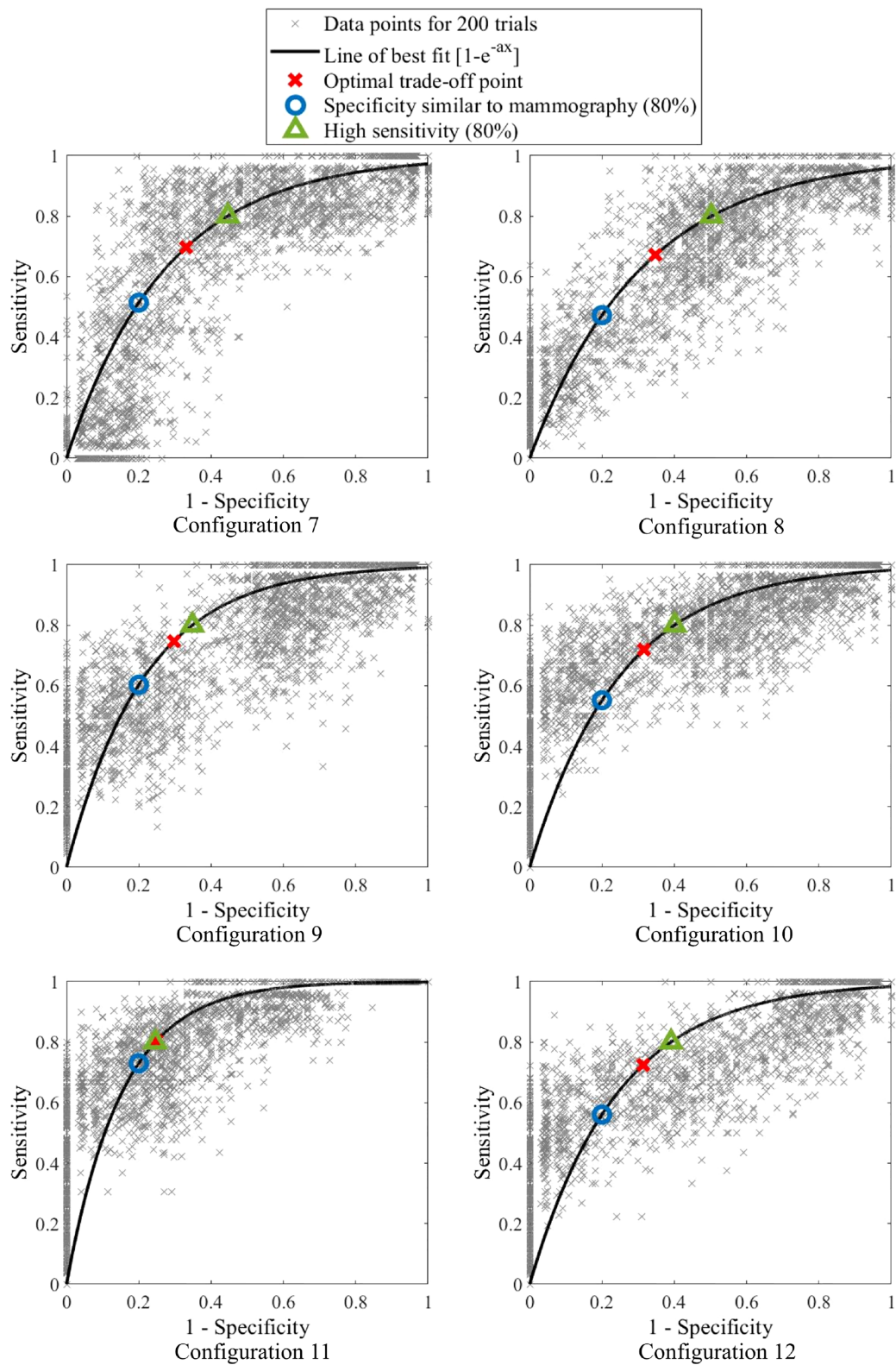


FIGURE 11
Bootstrapped ROC curves for breast configurations 7-12.

TABLE 3 Area under ROC curve (AUC), optimal performance point and assessment of diagnostic accuracy criteria for bootstrapping of each different breast configuration with bold values indicating the criteria is met.

Configuration	1	2	3	4	5	6	7	8	9	10	11	12
AUC	0.80	0.70	0.68	0.74	0.78	0.85	0.73	0.70	0.79	0.76	0.85	0.76
Optimal Sensitivity	76	67	66	71	74	81	70	67	75	72	80	72
Optimal Specificity	71	65	64	68	70	75	67	65	70	68	75	69
Sensitivity (with 80% specificity)	63	47	45	53	60	74	55	47	60	55	73	56
Specificity (with 80% sensitivity)	68	50	46	58	64	76	62	50	65	60	75	61

The bold values show configurations, which meet the diagnostic criteria.

TABLE 4 Assessing frequency composition method against diagnostic criteria in Table 2.

Criteria	Yes/No	Values
Unbiased diagnostic	Yes	
Full automation	Yes	
Ability to diagnose tumors down to 7mm	Yes	
Robust to varying breast sizes and densities	Yes	
AUC greater than mammography (>0.73)	Yes	0.85
Specificity 80%, sensitivity 60% or higher (1. similar to mammography)	Yes	74%
Sensitivity 80%, specificity 65% or higher (2. highly sensitive)	Yes	76%

(52, 61), and, as such, these false positive patients may have dense breasts, causing known issues for diagnosis using mammography, as dense tissue masks the presence of a tumor (63). Unfortunately, no follow up information is available with this data set regarding each patient's outcomes and consequent screenings. Test information is given as a one-off and, as such, we may never know this outcome.

Important to note is the right breast in patient P14R, which was correctly diagnosed as healthy in both configurations. This subject was originally diagnosed with cancer in mammography, which was later proven to be healthy tissue. Successful identification of this breast using this method shows an instance where DIET diagnostic capabilities were able to outperform mammography. This outcome helps prove implementation of DIET into the breast screening system could potentially improve overall diagnostic accuracy. Patient P6's right breast was also correctly identified as healthy in both configurations, despite having a non-malignant cyst, again showing the potential for DIET algorithms to distinguish between tumors and non-malignant lesions based on tissue stiffness.

With current limited clinical data, the method is primarily focused on the detection of tumor presence as a binary labelling problem. Detecting the exact location and depth of tumor for surgery and treatment purpose would require a much larger cohort of data to build its non-linear correlation to tissue properties and motion dynamics to avoid over-fitting issues. However, the current result did imply the frequency of interest for cancerous segments presented a notable contrast of response to healthy segments, which could be used to provide a preliminary estimation of location. Therefore, the benefit of the method is the comfort and ease of screening and the

automated results, which keeps running costs low, increases breast screening equity and encourages screening participation. While location around the breast segment could be achieved with this algorithm, further clinical breast exam, ultrasound, mammography or MRI would be recommended for confirming the exact location and depth.

Algorithm robustness and configuration selection

Figures 10, 11 show the bootstrapped ROC curves for all breast configurations following bootstrapping with 50 breasts selected with replacement and a repetition of 200 trials. Table 3 shows the AUC, optimal sensitivity and specificity, as well as sensitivity and specificity for criteria in Table 2 for each configuration. Bold values show configurations which meet the accuracy and AUC criteria.

Nine configurations (1, 4–7, 9–12) met the criteria for AUC over 0.73, which shows the algorithm is fairly robust to configuration selection, although some configurations are clearly more optimal. Four configurations (1, 6, 9, 11) met criteria for specificity greater than 65% when a highly sensitive (80%) diagnosis is achieved (Criteria #2). These configurations and an additional configuration 5 also met criteria for sensitivity at least 60% when specificity is similar to mammography (80%) (Criteria #1).

In general, the most optimal breast configurations occurred towards the top and bottom of the breast. The increased diagnostic quality of these segments might be attributed to the natural way the

breast hangs. Breast tissue structure is inhomogeneous and its complex structure and fibrous frame continually change with age due to effects, such as gravity (58, 64). The DIET set up may also distort surface motion in the natural hanging position due to pre-tensioning and pre-compression of surface tissue. It is possible using control segments at the top and bottom of the breast result in maximum surface tension, including the presence of suspensory (cooper's) ligaments (64) and, thus provide the truest steady state response for frequency analysis with minimum non-homogeneous tissue mechanics. Thus, these configurations result in optimal diagnosis for this frequency dependent diagnostic algorithm.

The optimal configuration (6) well exceeded performance criteria in Table 2 with optimal AUC at 0.85, optimal sensitivity and specificity of 81% and 75%. Sensitivity was 74% when specificity was similar to mammography (Criteria #1) and specificity was 76% when a highly sensitive diagnostic was achieved (Criteria #2). As mentioned, AUC of 0.85 is considered excellent (>0.8) (41), and not only well exceeds criteria (0.73), but exceeds all AUC values identified in studies on mammography (0.54–0.84) (12, 31, 40, 42–46).

As shown in Figures 6, 7, a combination of configuration 1 and 6 could result in perfect sensitivity at 100% and specificity of 69% exceeding criteria for a highly sensitive diagnostic ($>65\%$ in Criteria #2). This highly successful diagnostic outcome proves diagnostic efficacy using DIET can be achieved, supporting further research and investment in this technology.

Assessment against diagnostic criteria

All diagnostic criteria outlined in Table 2 were met or exceeded by the diagnostic method described in this paper. Table 4 shows optimal AUC well exceeds criteria at 0.85 (>0.73), and both accuracy criteria are well exceeded with Criteria #1 sensitivity at 74% ($>60\%$) and Criteria #2 specificity at 76% ($>65\%$). Furthermore, both 7 mm tumors were able to be correctly diagnosed in one of two configurations analysed, showing diagnosis of tumors, below the average tumor size detected by mammography at 10–14 mm (27, 65).

Limitations

The main limitation of this study is the limited clinical data available. This study presents results based 26 breasts analysed from 14 patients from a limited technical trial. Increased funding would enable more clinical trials and thus increased data, greatly improving validation of results and allowing for deep learning techniques to be utilised. However, the data in this cohort includes patients with a range of breast sizes and varying

tumor sizes and depths (Table 1), which demonstrate the success and potential of diagnostic algorithm across a varied cohort. Lack of displacement data for some patients at relevant input frequencies resulted in exclusion of one breast from each of Patient P4 and P13 in the analysis presented in this paper. All excluded breasts were healthy, so assessing the ability of this algorithm to diagnose cancerous breasts was not affected.

It should be noted, machine learning methods have been successfully applied for identification, diagnostic and analysis of medical data (66, 67). However, training the models of machine learning and deep neural networks normally require a very large cohort of labelled data in the medical field, which is not available in this case. Moreover, very detailed patient demographics and extra examinations might be needed to construct efficient input features to the training models, which may not necessarily be practical for a quick and equitable screening implementation such as *via* the DIET system. Therefore, a physics-based method and approach is considered to be more appropriate for this technology and at this time than machine learning methods, given the currently limited data for the DIET screening system. In future, such approaches could provide significant new diagnostic approaches.

Another limitation identified was the fitting of the ROC curve equation $y=1-e^{-ax}$ for some configurations. This equation was chosen at it is able to represent the linear (50:50 chance) and, with a very large exponent, the perfect square ROC curve, as well as all squares in between. However, it is restricted by the use of only one parameter (a), meaning there is a fixed relationship between sensitivity and specificity. Increasing the number of fitting parameters, such as $y=1-e^{-a(x-b)}+c$, would potentially over fit the data and may result in not meeting the 0 to 1 bounds of the ROC curve.

For instance when comparing configuration 2 in Figure 10 and configuration 8 in Figure 11, both have a similar bootstrapped ROC curve shape. Table 3 show identical AUC, optimal sensitivity and specificity and criteria values. Despite these similarities, the trend of the curve differs. Configuration 2 tends towards higher specificity, whereas configuration 8 appears more sensitive. These subtle differences are not captured by the fitting of equation $y=1-e^{-ax}$ with just one parameter. This issue is a limitation of this fit and development of this ROC curve model may more successfully capture the trade-off of sensitivity and specificity in some configurations. In general, the equation was successful in comparing configurations and most curves captured the general trend of points.

Conclusions

This paper presents a computationally efficient diagnostic algorithm, which meets identified criteria for comparable accuracy to mammography and the ability to provide a highly

sensitive diagnostic in breast screening. Three-dimensional plots showing response frequencies demonstrate how cancerous breasts exhibit higher and more varied frequencies of interest. An unbiased diagnostic was developed using a segmentation methodology, comparing second dominant frequencies in various breast segments, with similar frequencies expected in healthy breasts and more distinguishable differences indicating potential tumors. This method allowed for each breast to be diagnosed independently, removing issues of highly variable breast properties on diagnostic success.

Patient data at frequencies analysed was available for a total of 26 breasts (13 healthy and 13 cancerous) from 14 patients. An optimal breast configuration and diagnostic tolerance resulted in 85% sensitivity and 77% specificity. Using two configurations on either side of the breast demonstrated how sensitivity could be increased to 100% with only one additional false positive (specificity 69%), still meeting criteria for a highly sensitive diagnostic with manageable false positives. All diagnostic criteria were well exceeded showing potential for diagnosis using DIET to exceed diagnostic accuracy of mammography, including one breast correctly identified using this method, which was a false positive in mammography. ROC curve AUC exceeded all identified AUC values for mammography at 0.85 (0.54–0.84) and when specificity was similar to mammography (80%) sensitivity far exceeded it at 74% (>60%). This study provides an unbiased, fully automated diagnostic algorithm capable of detecting all tumors in this cohort, with manageable false positives, proving the diagnostic potential of the DIET technology, as a breast screening tool with many benefits.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Clinical data can be requested from Tiromedical, NZ. Requests to access these datasets should be directed to jessica.fitzjohn@pg.canterbury.ac.nz.

References

- van den Ende C, Oordt-Speets AM, Vrolijk H, van Agt HME. Benefits and harms of breast cancer screening with mammography in women aged 40–49 years: A systematic review: Breast cancer screening in women aged 40–49 years. *Int J Cancer* (2017) 141:1295–306. doi: 10.1002/ijc.30794
- Lotz T, Muller N, Hann CE, Chase JG. Minimal elastographic modeling of breast cancer for model based tumour detection in a digital image elasto tomography (DIET) system. *Proc SPIE Med Imaging Conf* (2011) 7963:1–6. doi: 10.1117/12.878315
- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: A Cancer J Clin* (2011) 61:69–90. doi: 10.3322/caac.20107
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
- Ahmad A. *Breast cancer metastasis and drug resistance: Challenges and progress, advances in experimental medicine and biology* Vol. 1152. . Cham: Springer International Publishing (2019). doi: 10.1007/978-3-030-20301-6
- Heywang-Köbrunner SH, Hacker A, Sedlacek S. Advantages and disadvantages of mammography screening. *Breast Care* (2011) 6:2–2. doi: 10.1159/000329005
- Loberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res* (2015) 17:1–12. doi: 10.1186/s13058-015-0525-z
- Hollingsworth AB. Redefining the sensitivity of screening mammography: A review. *Am J Surg* (2019) 218:411–8. doi: 10.1016/j.amjsurg.2019.01.039
- Moshina N, Sebuodegard S, Lee CI, Akslen LA, Tsuruda KM, Elmore JG, et al. Automated volumetric analysis of mammographic density in a screening setting: Worse outcomes for women with dense breasts. *Radiology* (2018) 288:343–52. doi: 10.1148/radiol.2018172972
- Berg WA. Tailored supplemental screening for breast cancer: What now and what next? *Am J Roentgenol* (2009) 192:390–9. doi: 10.2214/AJR.08.1706
- Faruk T, Islam MK, Arefin S, Haq MZ. The journey of elastography: Background, current status, and future possibilities in breast cancer diagnosis. *Clin Breast Cancer* (2015) 15:313–24. doi: 10.1016/j.clbc.2015.01.002

Ethics statement

The studies involving human participants were reviewed and approved by New Zealand Health and disability ethics committee (URA/08/08/059). The patients/participants provided their written informed consent to participate in this study.

Author contributions

JF did the conceptualisation, analysis and writing, CZ and JGC provided conceptualisation, aided analysis and revised writing. All authors contributed to the article and approved the submitted version.

Funding

The authors acknowledge UC Doctoral Scholarship funding.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

12. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al. Diagnostic accuracy of digital versus film mammography: Exploratory analysis of selected population subgroups in DMIST. *Radiology* (2008) 246:376–83. doi: 10.1148/radiol.2461070200
13. Margolese R. Screening mammography in young women: A different perspective. *Lancet* (1996) 347:881–2. doi: 10.1016/S0140-6736(96)91352-5
14. Williams SM, Kaplan PA, Petersen JC, Lieberman RP. Mammography in women under age 30: Is there clinical benefit? *Radiology* (1986) 161:49–51. doi: 10.1148/radiology.161.1.3763885
15. Barlow WE. Performance of diagnostic mammography for women with signs or symptoms of breast cancer. *CancerSpectrum Knowl Environ* (2002) 94:1151–9. doi: 10.1093/jnci/94.15.1151
16. Houssami N, Ciatto S, Martinelli F, Bonardi R, Duffy S. Early detection of second breast cancers improves prognosis in breast cancer survivors. *Ann Oncol* (2009) 20:1505–10. doi: 10.1093/annonc/mdp037
17. Humphrey LL, Helfand M, Chan BKS, Woolf SH. Breast cancer screening: A summary of the evidence for the U.S. preventive services task force. *Ann Internal Med* (2002) 137:347–60. doi: 10.7326/0003-4819-137-5part1-200209030-00012
18. Lekanidi K, Dilks P, Suaris T, Kennett S, Purushothaman H. Breast screening: What can the interval cancer review teach us? are we perhaps being a bit too hard on ourselves? *Eur J Radiol* (2017) 94:13–5. doi: 10.1016/j.ejrad.2017.07.005
19. Peters A, Milsant A, Rouze J, Ray L, Chase JG, Houten EEWV. Digital image-based elasto-tomography: Proof of concept studies for surface based mechanical property reconstruction. *JSME International Journal Ser C Mech Syst Mach Elem Manuf* (2004) 47:1117–23. doi: 10.1299/jsmec.47.1117
20. Peters A, Chase JG, Van Houten EE. Digital image elasto-tomography: Mechanical property estimation of silicone phantoms. *Med Biol Eng Comput* (2008) 46:205–12. doi: 10.1007/s11517-007-0275-x
21. Peters A, Wortmann S, Elliott R, Staiger M, Chase JG, Houten EV. Digital image-based elasto-tomography: First experiments in surface based mechanical property estimation of gelatine phantoms. *JSME Int J Ser C Mech Syst Mach Elem Manuf* (2005) 48:562–9. doi: 10.1299/jsmec.48.562
22. Kashif AS, Lotz TF, Heeren AMW, Chase JG. Separate modal analysis for tumor detection with a digital image elasto tomography (DIET) breast cancer screening system: Separate modal analysis for tumor detection. *Med Phys* (2013) 40:113503. doi: 10.1118/1.4826168
23. Kashif AS, Lotz TF, McGarry MD, Pattison AJ, Chase JG. Silicone breast phantoms for elastographic imaging evaluation. *Med Phys* (2013) 40:063503. doi: 10.1118/1.4805096
24. Samani A, Zubovits J, Plewes D. Elastic moduli of normal and pathological human breast tissues: An inversion-technique-based investigation of 169 samples. *Phys Med Biol* (2007) 52:1565–76. doi: 10.1088/0031-9155/52/6/002
25. Kopans DB. *Breast imaging*. (Philadelphia: Lippincott Williams & Wilkins) (1998).
26. Zhou C, Chase JG, Ismail H, Signal MK, Haggars M, Rodgers GW, et al. Silicone phantom validation of breast cancer tumor detection using nominal stiffness identification in digital imaging elasto-tomography (DIET). *Biomed Signal Process Control* (2018) 39:435–47. doi: 10.1016/j.bspc.2017.08.022
27. Zhou C, Hainsworth B, Sydney M, Lee M, Ormsby Z, Haggars M, et al. Structural health monitoring of tissue mechanics for non-invasive diagnosis of breast cancer. *Automatisierungstechnik* (2018) 66:1037–50. doi: 10.1515/auto-2018-0065
28. Ismail HM, Pretty CG, Signal MK, Haggars M, Chase JG. Finite element modelling and validation for breast cancer detection using digital image elasto-tomography. *Med Biol Eng Comput* (2018) 56:1715–29. doi: 10.1007/s11517-018-1804-5
29. Ismail HM, Pretty CG, Signal MK, Haggars M, Zhou C, Chase JG. Mechanical behaviour of tissue mimicking breast phantom materials. *Biomed Phys Eng Express* (2017) 3:045010. doi: 10.1088/2057-1976/aa7992
30. Fitzjohn JL, Zhou C, Chase JG, Ormsby Z, Haggars M. Modeling viscous damping in actuated breast tissue to provide diagnostic insight for breast cancer: A proof-of-concept analysis. *Med Phys* (2021) 48:4978–92. doi: 10.1002/mp.15054
31. Roganovic D, Djilas D, Vujnovic S, Pavic D, Stojanov D, Breast MRI. Digital mammography and breast tomosynthesis: Comparison of three methods for early detection of breast cancer. *Bosnian J Basic Med Sci* (2015) 14(4):64–8. doi: 10.17305/bjbm.2015.616
32. Brem RF, Taba' r L, Duffy SW, Inciardi MF, Guingrich JA, Hashimoto BE, et al. Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue. *SonoInsight Study Radiol* (2015) 274:663–73. doi: 10.1148/radiol.14132832
33. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast us and evaluation of factors that influence them: An analysis of 27,825 patient evaluations 1. *Radiology* (2002) 225:165–75. doi: 10.1148/radiol.2251011667
34. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *Jama* (2012) 307:1394–404. doi: 10.1001/jama.2012.388
35. de Oliveira Pereira R, da Luz LA, Chagas DC, Amorim JR, Nery-Júnior E, Alves ACBR, et al. Evaluation of the accuracy of mammography, ultrasound and magnetic resonance imaging in suspect breast lesions. *Clinics* (2020) 75:e1805. doi: 10.6061/clinics/2020/e1805
36. Devolli-Disha E, Manxhuka-Kërliu S, Ymeri H, Kutllövci A. Comparative accuracy of mammography and ultrasound in women with breast symptoms according to age and breast density. *Bosnian J Basic Med Sci* (2009) 9:131–6. doi: 10.17305/bjbm.2009.2832
37. Skaane P, Bandos AI, Niklason LT, Sebuodegård S, Osterås BH, Gullien R, et al. DigitalMammography versus DigitalMammography plus tomosynthesis in breast cancer screening. *Oslo Tomosynthesis Screen Trial Radiol* (2019) 291:23–30. doi: 10.1148/radiol.2019182394
38. Sardanelli F, Podo F. Breast MR imaging in women at high-risk of breast cancer: is something changing in early breast cancer detection? *Eur Radiol* (2007) 17:873–87. doi: 10.1007/s00330-006-0389-9
39. Egan R, Egan K. Detection of breast carcinoma: Comparison of automated water-path whole-breast sonography, mammography, and physical examination. *Am J Roentgenol* (1984) 143:493–7. doi: 10.2214/ajr.143.3.493
40. Giger ML, Inciardi MF, Edwards A, Papaioannou J, Drukker K, Jiang Y, et al. Automated breast ultrasound in breast cancer screening of WomenWith dense breasts: Reader study of mammography-negative and mammography-positive cancers. *Am J Roentgenol* (2016) 206:1341–50. doi: 10.2214/AJR.15.15367
41. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* (2010) 5:1315–6. doi: 10.1097/JTO.0b013e3181ec173d
42. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *New Engl J Med* (2005) 353:1773–83. doi: 10.1056/NEJMoa052911
43. Marini C, Traino C, Cilotti A, Roncella M, Campori G, Bartolozzi C. Differentiation of benign and malignant breast microcalcifications: Mammography versus mammography-sonography combination. *La Radiol Med* (2003) 105:17–26.
44. Hadjiiski L, Chan HP, Sahiner B, Helvie MA, Roubidoux MA, Blane C, et al. Improvement in radiologists' characterization of malignant and benign breast masses on serial mammograms with computer-aided diagnosis: An ROC study. *Radiology* (2004) 233:255–65. doi: 10.1148/radiol.2331030432
45. Boita J, van Engen RE, Mackenzie A, Tingberg A, Bosmans H, Bolejko A, et al. Validation of a candidate instrument to assess image quality in digital mammography using ROC analysis. *Eur J Radiol* (2021) 139:109686. doi: 10.1016/j.ejrad.2021.109686
46. Rangayyan R, Shen L, Desautels J, Bryant H, Terry T, Shen Y, et al. Improvement of sensitivity of breast cancer diagnosis with adaptive neighborhood contrast enhancement of mammograms. *IEEE Trans Inf Technol Biomed* (1997) 1(3):161–70. doi: 10.1109/4233.654859
47. Ling SJ, Sanny J, Moebis W. University physics volume. (2021) 1:979.
48. Cochran W, Cooley J, Favini D, Helms H, Kaenel R, Lang W, et al. What is the fast Fourier transform? *Proc IEEE* (1967) 55:1664–74. doi: 10.1109/PROC.1967.5957
49. Li H, Zhang S, Wang Q, Zhu R. Clinical value of mammography in diagnosis and identification of breast mass. *Pakistan J Med Sci* (2016) 32:1020–25. doi: 10.12669/pjms.324.9384
50. Johnson M. What are the most common breast shapes. (2019) 1(2):1.
51. Dickson JL, Gunn CA, Chase JG. Title : Humans are horribly variable. *Int J Clin Med Imaging* (2014) 1:1. doi: 10.4172/ijcmi.1000142
52. Gusterson BA, Stein T. Human breast development. *Semin Cell Dev Biol* (2012) 23:567–73. doi: 10.1016/j.semcdb.2012.03.013
53. Anthony K. Breast asymmetry. (2018).
54. Taroni P, Quarto G, Pifferi A, Abbate F, Balestreri N, Menna S, et al. Breast tissue composition and its dependence on demographic risk factors for breast cancer: Non-invasive assessment by time domain diffuse optical spectroscopy. *PloS One* (2015) 10:e0128941. doi: 10.1371/journal.pone.0128941
55. Vazirinejad R, Darakhshan S, Esmaili A, Hadadian S. The effect of maternal breast variations on neonatal weight gain in the first seven days of life. *Int Breastfeed J* (2009) 4:13. doi: 10.1186/1746-4358-4-13
56. Samani A, Bishop J, Luginbuhl C, Plewes DB. Measuring the elastic modulus of ex vivo small tissue samples. *Phys Med Biol* (2003) 48:2183–98. doi: 10.1088/0031-9155/48/14/310
57. Krouskop TA, Wheeler TM, Kallel F, Garra BS, Hall T. Elastic moduli of breast and prostate tissues under compression. *Ultrason Imaging* (1998) 20:260–74. doi: 10.1177/016173469802000403
58. Ramião NG, Martins PS, Rynkevicius R, Fernandes AA, Barroso M, Santos DC. Biomechanical properties of breast tissue, a state-of-the-art review. *Biomech Model Mechanobiol* (2016) 15:1307–23. doi: 10.1007/s10237-016-0763-8

59. Elmore JG, Barton MB, Mocer VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *New Engl J Med* (1998) 338:1089–96. doi: 10.1056/NEJM199804163381601
60. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI, et al. Revision of the American joint committee on cancer staging system for breast cancer. *J Clin Oncol* (2002) 20:3628–36. doi: 10.1200/JCO.2002.02.026
61. Faridah Y. Digital versus screen film mammography: A clinical comparison. *Biomed Imaging Interv J* (2008) 4:e31. doi: 10.2349/bij.4.4.e31
62. Mrose H. Small breast sizes are also at risk of cancer. (2020).
63. Destounis S, Arieno A, Morgan R, Roberts C, Chan A. Qualitative versus quantitative mammographic breast density assessment: Applications for the US and abroad. *Diagnostics* (2017) 7:30. doi: 10.3390/diagnostics7020030
64. Gefen A, Dilmoney B. Mechanics of the normal woman's breast. *Technol Health Care* (2007) 15:259–71. doi: 10.3233/THC-2007-15404
65. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *Jama* (2008) 299:2151–63. doi: 10.1001/jama.299.18.2151
66. Iwendi C, Huescas CGY, Chakraborty C, Mohan S. COVID-19 health analysis and prediction using machine learning algorithms for Mexico and Brazil patients. *J Exp Theor Artif Intell* (2022) 1–21. doi: 10.1080/0952813X.2022.2058097
67. Vijayalakshmi S, John A, Senthilkumar R, Mohan S, Bhattacharya S, Kaluri R, et al. Multi-modal prediction of breast cancer using particle swarm optimization with non-dominating sorting. *Int J Distr Sensor Networks* (2020) 16:155014772097150. doi: 10.1177/1550147720971505



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

Yueh Z. Lee,
University of North Carolina,
United States
Moreno Zanardo,
University of Milan, Italy

*CORRESPONDENCE

Natalia Siminiak
nsiminiak@gmail.com

SPECIALTY SECTION

This article was submitted to
Breast Cancer,
a section of the journal
Frontiers in Oncology

RECEIVED 11 May 2022

ACCEPTED 04 November 2022

PUBLISHED 24 November 2022

CITATION

Siminiak N, Pasiuk-Czepczyńska A,
Godlewska A, Wojtyś P, Olejnik M,
Michalak J, Nowaczyk P, Gajdzis P,
Godlewski D, Ruchała M and
Czepczyński R (2022) Are contrast
enhanced mammography and digital
breast tomosynthesis equally effective
in diagnosing patients recalled from
breast cancer screening?
Front. Oncol. 12:941312.
doi: 10.3389/fonc.2022.941312

COPYRIGHT

© 2022 Siminiak, Pasiuk-Czepczyńska,
Godlewska, Wojtyś, Olejnik, Michalak,
Nowaczyk, Gajdzis, Godlewski, Ruchała
and Czepczyński. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Are contrast enhanced mammography and digital breast tomosynthesis equally effective in diagnosing patients recalled from breast cancer screening?

Natalia Siminiak^{1*}, Anna Pasiuk-Czepczyńska²,
Antonina Godlewska¹, Piotr Wojtyś², Magdalena Olejnik²,
Joanna Michalak², Piotr Nowaczyk³, Paweł Gajdzis²,
Dariusz Godlewski², Marek Ruchała¹ and Rafał Czepczyński¹

¹Department of Endocrinology, Metabolism and Internal Diseases, Poznan University of Medical
Sciences, Poznań, Poland, ²Cancer Prevention and Epidemiology Center, Poznań, Poland, ³Breast
Surgical Oncology Department, Greater Poland Cancer Center, Poznań, Poland

Purpose: Full-field digital mammography (FFDM) is widely used in breast cancer screening. However, to improve cancer detection rates, new diagnostic tools have been introduced. Contrast enhanced mammography (CEM) and digital breast tomosynthesis (DBT) are used in the diagnostic setting, however their accuracies need to be compared. The aim of the study was to evaluate the diagnostic performance of CEM and DBT in women recalled from breast cancer screening program.

Methods: The study included 402 consecutive patients recalled from breast cancer screening program, who were randomized into two groups, to undergo either CEM (202 patients) or DBT (200 patients). All visible lesions were evaluated and each suspicious lesion was histopathologically verified.

Results: CEM detected 230 lesions; 119 were classified as benign and 111 as suspicious or malignant, whereas DBT identified 209 lesions; 105 were classified as benign and 104 as suspicious or malignant. In comparison to histopathology, CEM correctly detected cancer in 43 out of 44 cases, and DBT in all 33 cases, while FFDM identified 15 and 18 neoplastic lesions in two groups, respectively. CEM presented with 97% sensitivity, 63% specificity, 70% accuracy, 38% PPV and 99% NPV, while DBT showed 100% sensitivity, 60% specificity,

32%, PPV, 100% NPV and 66% accuracy. The CEM's AUC was 0.97 and DBT's 0.99. The ROC curve analysis proved a significant ($p < 0.000001$) advantage of both CEM and DBT over FFDM, however, there was no significant difference between CEM and DBT diagnostic accuracies ($p = 0.23$).

Conclusions: In this randomized, prospective study CEM and DBT show similar diagnostic accuracy.

KEYWORDS

breast cancer, contrast enhanced mammography, digital breast tomography (DBT), mammography, imaging modalities

Introduction

Mammography (Full Field Digital Mammography, FFDM) is the most common and widely available imaging modality for breast cancer diagnosis. FFDM is used in routine screening, in symptomatic patients, and often in monitoring primary systemic breast cancer therapy, as well as in the follow-up after treatment (1). In 2011, Food and Drug Administration (FDA) approved two new modalities: contrast enhanced mammography (CEM) and digital breast tomosynthesis (DBT) for clinical use as adjuncts to mammography (2, 3).

CEM is a promising diagnostic technique, especially in women with dense parenchymal tissue (4). By administering an intravenous contrast agent, CEM adds new physiological information to the morphological data provided by FFDM and improves the parameters of breast cancer detection (5). In a recent meta-analysis, performed on 60 studies including 11049 examinations, Cozzi et al. reported CEM sensitivity of 95% and specificity of 81% (6). The negative predictive value (NPV) of CEM reaches 100%, and the positive predictive value (PPV) varies from 93 to 97% (7, 8). In contrast, FFDM's sensitivity remains at the level of 86–89% in low density, fatty breasts and around 62–68% in dense breasts (1).

Digital breast tomosynthesis (DBT) is a subtype of mammography, which uses the same X-ray source and creates multiple 2D images to obtain a 3D breast reconstruction (3, 9). DBT outperforms conventional mammography mostly due to the ability to evaluate overlapping breast tissue and visualize tumors, which are not visible on FFDM. Researchers estimated that the sensitivity of DBT is higher than that of FFDM, around 81.1% vs 60.4% respectively (10).

Both CEM and DBT are relatively recent/new imaging techniques and are constantly being developed. Imaging protocols, radiation dose, or image acquisition are being improved for better sensitivity, specificity, but also patient safety (11). In order to introduce new techniques to routine screening and diagnosis, it is important to assess their diagnostic value.

This paper presents novel data on CEM and DBT diagnostic efficiency. The purpose of the study was to evaluate the diagnostic performance of CEM and DBT in a cohort of women recalled from the national mammography screening program.

Material and methods

Patients

The study is a prospective, randomized trial, approved by the regional Medical Bioethical Committee of Greater Poland Medical Chamber. Patients involved in the study were invited to participate after being recalled from routine breast cancer screening. Patients aged from 50 to 69 years, recalled from screening, were randomized with allocation ratio 1:1 into two groups in which either CEM of both breasts or DBT of suspected breasts were performed. Patients were allocated to each group using computer-assisted randomization. The study was performed at the Cancer Prevention and Epidemiology Center in Poznań between July 2020 and September 2021. The written consent of each patient was obligatory for participation. Exclusion criteria were: signs and symptoms of breast cancer, previous breast surgery including implants, $eGFR < 30 \text{ ml/min} \times 1.73 \text{ m}^2$, history of allergy to iodinated contrast agents, lack of patient's consent to participate in the trial.

Full field digital mammography

FFDM was performed in all of the patients in a screening setting. All examinations were performed at the Center of Cancer Prevention and Epidemiology in Poznań, either in mobile (Mammomat Fusion, Siemens Healthcare, Germany) or stationary mammography units (Mammomat Inspiration Prime and Mammomat Revelation, Siemens).

All of the devices placed in mobile units were equipped with a flat panel detector with a cesium iodide scintillator, field size 230x300 mm, a 2790 × 3580 image matrix with a detector elements pitch of 83 µm.

Contrast enhanced mammography

CEM was performed in recalled patients 2-3 weeks after the screening mammograms. All of the contrast enhanced examinations were performed by means of a stationary mammography unit (Mammomat Revelation, Siemens). It consisted of a current FFDM system using a flat panel detector (amorphous selenium (aSe)) with a field size 299x238 mm, a 2800x3518 image matrix with a detector elements pitch of 85 µm and specific software and hardware for rapid acquisition and processing of dual-energy images.

Low energy images were acquired 2 minutes after intravenous iopromide 370 (Ultravist 370, Bayer Healthcare, Berlin, Germany) administration. It was a one-shot injection using a power injector Dual Shot alpha 7 (Nemoto) of 1,5 ml/kg at a rate of 3 ml/s with a 30 ml saline bolus chaser. CEM examinations began with a mediolateral oblique (MLO) view of the breast without suspicious lesion following the breast with the suspicious lesion in order to secure more contrast uptake. The pair of low and high energy images (LE and HE) was performed on each view (MLO and cranio-caudal (CC)). Then recombined images (RC) were generated in order to visualize enhancing lesions and eliminate unenhancing background tissue. Tungsten anode material was used for all acquisitions, with a rhodium filter with kVp ranging from 25 to 32 used for low energy acquisitions similar to those in FFDM. The paired high-energy images were acquired at a 49 kVp titanium (Ti) filter in the X-ray beam to produce an X-ray spectrum above the K-edge of iodine (33.2 KeV), to increase the visibility of low concentrations of iodine (12). The entrance dose varied from 1.26 to 12.07 mGy, depending on the thickness of the breast (10 - 82 mm) and tissue composition.

All of the lesions visible in CEM were histopathologically verified, either by core needle biopsy or vacuum-assisted biopsy under ultrasound or mammography guidance.

Digital breast tomosynthesis

DBT procedures were performed 2-3 weeks after screening mammograms, using a digital mammography Mammomat Inspiration device (Siemens). It consisted of a current FFDM system using a flat panel amorphous selenium (aSe) detector with a field size of 299 x 238 mm, a 2800 × 3518 image matrix with a detector elements pitch of 85 µm, and specific software and hardware for rapid acquisition and processing of tomographic images.

The tube swivel range was 50° (+/-25°) with 25 projections and with a 1mm distance between reconstructed slices. Two views of the breast with the suspicious lesion(s) were performed, mostly mediolateral oblique (MLO) and CC. The entrance dose varied from 2,46 to 14,92 mGy, depending on the thickness of the breast (23 - 86 mm) and tissue composition.

All of the lesions visible on DBT were biopsied and histopathologically proven.

Image evaluation

Vue PACS review workstation (Carestream) was used for image analysis. Two radiologists with 8 and 15 years of experience in breast imaging evaluated the recalled patients either with DBT or CEM. In practice, each CEM and DBT examination was evaluated independently by each radiologist and the results of their assessment were recorded. In case of discordance, the examination was reviewed by both readers and the consensus was recorded as the final result.

The two radiologists were concordant in 179 cases

All the mammograms of recalled patients were evaluated according to the ACR Breast Imaging Reporting and Data System[®] (ACR - BIRADS (13, 14)). All the recalled patients' baseline mammograms were classified as BIRADS 0 - demanding further evaluation to be classified into adequate BIRADS category (1, 2, 3, 4 or 5), BIRADS 4 - suspicious, BIRADS 5 - malignant.

Subsequently, lesions of patients from CEM and DBT groups were reported in concordance with the CEM supplement to ACR BIRADS[®] Mammography 2013 (14) and categorized into the following BIRADS groups: BIRADS 1 - negative, BIRADS 2 - benign, BIRADS 3 - probably benign, BIRADS 4 - suspicious, BIRADS 5 - highly suggestive of malignancy.

Patients with BIRADS 1 are being followed up with CEM or DBT respectively, performed after 12 and 24 months. All the lesions classified above BIRADS 1 were biopsied. BIRADS classification outcome was compared to histopathology results. Finally, on the basis of radio-pathological concordance, lesions were classified as:

- 1) true positive (BIRADS \geq 4 and biopsy-proven cancer),
- 2) false positive (BIRADS \geq 4 and biopsy-proven benign lesion),
- 3) false negative (BIRADS \leq 3 and biopsy-proven cancer),
- 4) true negative (BIRADS \leq 3 and biopsy-proven benign lesion).

All in all, the results of each patient group were compared (Tables 1, 2).

TABLE 1 Distribution of lesions visible in CEM.

Lesions detected by CEM

Histopathological findings	Number of lesions detected
Benign Lesions	61
fibrocystic changes	19
fibroadenoma	10
intraductal papilloma	6
fragments of atrophic breast tissue	4
fragments of breast tissue	4
sclerosing adenosis	4
intraductal papilloma with UDH	3
intramammary lymph node	3
fibrocystic changes with UDH	3
inflammation due to hidradenitis suppurativa (acne inversa)	2
fibroadenoma with usual ductal hyperplasia (UDH)	1
fibrocystic changes with sclerosing adenosis	1
columnar cell changes	1
Suspicious (B3) Lesions	6
atypical lobular hyperplasia	3
atypical ductal hyperplasia (ADH)	2
breast papilloma with ADH	1
Cancers	44
invasive carcinoma of no special type (NST)	30
invasive lobular carcinoma	7
invasive carcinoma of no special type (NST) with ductal carcinoma <i>in situ</i> (DCIS)	3
invasive carcinoma of no special type (NST), partially mucinous, with DCIS	1
invasive carcinoma of no special type (NST), partially micropapillary	1
invasive lobular carcinoma with lobular carcinoma <i>in situ</i>	1
invasive lobular carcinoma with invasive carcinoma NST and lobular carcinoma <i>in situ</i> (LCIS)	1

Statistical analysis

The calculations, including the sample size, were made using Statistica 13 by TIBCO and PQStat by PQStat Software. The level of significance was $\alpha = 0.05$. The result was considered statistically significant when $p < \alpha$. The normality of the distribution of variables was tested with the Shapiro-Wilk test. In order to compare the variables between the two groups, the Mann-Whitney test was calculated because of non-compliance with the normal distribution. The correlation between categorical variables was calculated using the χ^2 test of independence or the Fisher-Freeman-Halton test. The

compliance of the methods of assessing the occurrence of neoplasms was tested by calculating Cohen's kappa coefficient of concordance and determining their significance using the Z test. The Fleiss kappa coefficient of concordance was calculated to test the consistency of all 3 methods (mammography, CEM/DBT, and biopsy) simultaneously. Additionally, the sensitivity, specificity, PPV, and NPV were determined with 95% confidence intervals. ROC analysis was performed to calculate the optimal cut-off point for BIRADS. The area under the curves (AUC) with 95% confidence intervals was determined using the non-parametric DeLong method. The optimal cut-off point was established using the Youden Index. Sensitivity and specificity were determined for the selected cut-off. The determined areas under the curve were compared with each other using the Z statistics.

Results

402 consecutive patients recalled from the national breast cancer screening program were included in the study. The sample size was sufficient for analysis according to the sample size power calculation (minimal sample size 151 subjects). One half of the patients (200) underwent FFDM in the screening setting followed by the CEM examination, whereas 202 patients underwent FFDM followed by the DBT examination. The two radiologists, who evaluated the images, were concordant with their diagnosis in 179 CEM cases (89.5%) and the consensus had to be achieved in the remaining 21 discordant results (10.5%). In case of the DBT, the two readers were concordant in 194 patients (96.0%) and discordant in 8 cases (4.0%).

CEM group

As detailed in the study flowchart (Figure 1), CEM indicated 230 lesions; 119 of them (52%) were described as benign (BIRADS 1 or 2) and 111 (48%) as suspicious or malignant (BIRADS ≥ 4). Histopathology examination confirmed cancer in 44 lesions. CEM was true positive in 43 cases, true negative in 118 cases, false positive in 68 cases, and there was 1 false-negative case.

FFDM, performed in CEM subgroup patients, indicated 205 lesions, where 171 were classified as benign (BIRADS = 1 or 2) and 34 were described as suspicious or malignant (BIRADS ≥ 4). FFDM was true positive in 15 cases and true negative in 148 cases, however, it was false negative in 23 cases and false positive in 19 patients.

CEM presented sensitivity of 97%, specificity 63%, accuracy of 70%, PPV of 38%, NPV of 99%, whereas FFDM in this

TABLE 2 Distribution of lesions visible in DBT.

Lesions detected by DBT

Histopathological findings	Number of lesions detected
Benign Lesions	67
fibrocystic changes	32
fibroadenoma	8
usual ductal hyperplasia (UDH)	5
fibrocystic changes with UDH	4
fibrocystic changes with sclerosing adenosis and microcalcifications	3
intramammary lymph node	3
sclerosing adenosis + simple adenosis	3
fibroadenoma with microcalcifications	2
fibroadenomatous like changes	2
fibrocystic changes with UDH and apocrine hyperplasia	1
hamartoma	1
fragments of breast tissue	1
fragments of breast tissue with signs of chronic inflammation	1
nodular adenosis	1
Suspicious (B3) Lesions	4
atypical ductal hyperplasia (ADH)	3
breast papilloma with ADH	1
Cancers	33
invasive carcinoma of no special type (NST)	18
invasive carcinoma of no special type (NST) with ductal carcinoma <i>in situ</i> (DCIS)	5
ductal carcinoma <i>in situ</i> (DCIS)	4
invasive lobular carcinoma	4
mucinous carcinoma	1
invasive carcinoma of no special type (NST), partially micropapillary	1

subgroup showed sensitivity of 40%, specificity 87%, PPV 44%, NPV 87% and accuracy 80% (Table 3).

In summary, CEM indicated 43 cancer lesions, whereas FFDM identified cancer in 15 cases.

The Kappa analysis confirmed a fair concordance level (Kappa = 0.29) between FFDM and biopsy, as well as between CEM and biopsy (Kappa = 0.39).

The ROC curves, based on BI-RADS classification, showed significant differences between CEM and FFDM examinations ($p < 0.000001$). CEM presented with AUC 0.97, while FFDM with AUC 0.65 (Figure 2).

DBT subgroup

DBT indicated 209 lesions, where 105 cases were described as benign (BIRADS = 1 or 2) and 104 lesions as suspicious or

malignant (BIRADS ≥ 4). Histopathology confirmed cancer in 33 tumors. DBT was true positive in 33 cases, true negative in 105 cases, and false-positive in 71 cases. DBT presented sensitivity of 100%, specificity of 60%, PPV 32%, NPV 100% and accuracy 66% (Table 3).

FFDM, performed prior to DBT, indicated 207 lesions, where 173 cases were described as benign (BIRADS = 1), and 34 cases as suspicious or malignant (BIRADS ≥ 4). FFDM was true positive in 18 cases and true negative in 157 cases, however, it was false negative in 18 cases and false positive in 16 cases. FFDM showed sensitivity of 55%, specificity of 91%, PPV 53%, NPV 91% and accuracy 85%.

In summary, DBT indicated 33 lesions as malignant and all of them were confirmed as cancers in the biopsy, whereas FFDM indicated cancer in 18 of these cases.

The Kappa analysis indicated a fair concordance level (Kappa = 0.32) between DBT and biopsy. However, a moderate concordance level (Kappa = 0.45) was established between FFDM and biopsy.

The ROC curves, based on BI-RADS classification, showed significant differences between DBT and mammography examinations ($p < 0.000001$). DBT presented with an AUC of 0.99, while mammography with AUC of 0.74 (Figure 3).

The comparison between CEM and DBT groups

The groups did not present any differences in the age of patients (average 59 years) as well as in the size of lesions (average size 12.3 mm in the DBT group and 11.8 mm in the CEM group).

The cancer detection rate was similar in both modalities ($p = 0.8$) and the percentage of diagnosed cancer cases (histopathological confirmation) in DBT and CEM were similar ($p = 0.35$).

The ROC curves based on BI-RADS classifications for DBT and CEM are located similarly, demonstrating their similar diagnostic abilities. There were no significant differences between ROC curve areas ($p = 0.23$) (Figure 4).

Discussion

Breast cancer is the most often diagnosed malignancy and the first cause of cancer death among women. The International Agency for Research on Cancer (IARC) reported 2.26 million cases of breast cancer in 2020 worldwide (16). Early detection continues to be the key to a better prognosis and higher survival rate. FFDM remains the gold standard in breast cancer screening, however novel diagnostic tools, which may improve its accuracy (CEM, DBT, magnetic resonance), have been introduced. To the best of our knowledge, this is the first

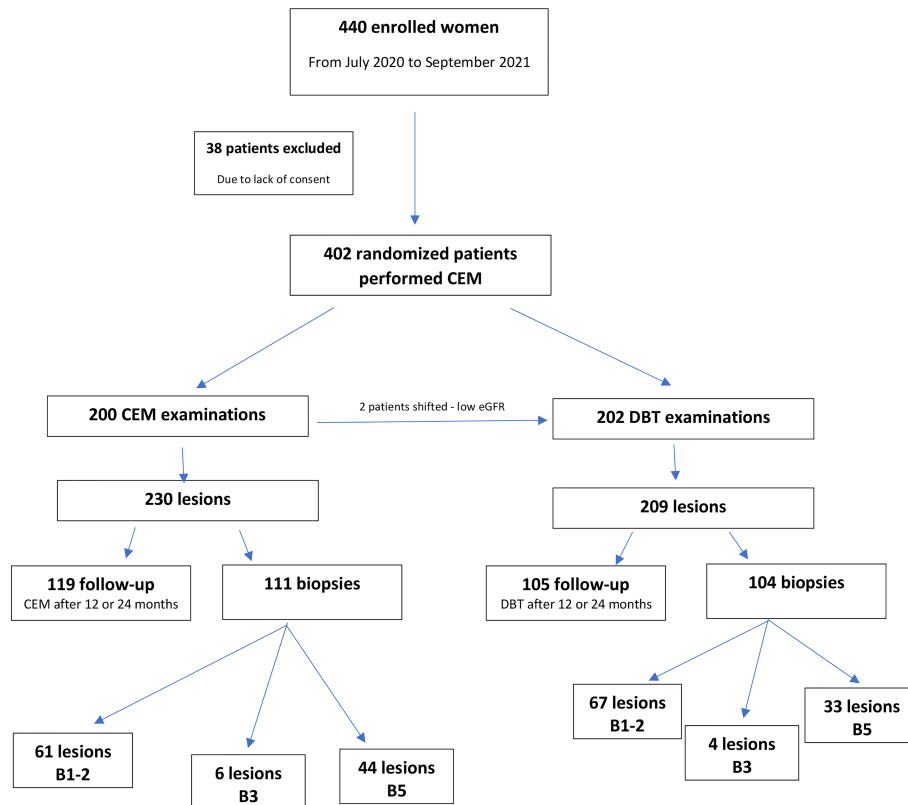


FIGURE 1

Study flowchart presenting distribution of patients and lesions due to the histopathology results estimated by core needle biopsy or vacuum-assisted biopsy under ultrasound or mammography guidance (15).

randomized, prospective study comparing CEM and DBT performance in patients recalled from breast cancer screening.

Studies have reported that both CEM and DBT have shown high sensitivity, especially with dense breasts, as compared to FFDM (17). In our study, the distribution of breast density patterns was similar in both groups. The analysis of our study

cohort divided into four density patterns (A, B, C, and D according to ACR) will be performed in the further stage of our project.

The analysis confirmed high sensitivity (97%) and specificity (63%) of CEM, similar to those previously reported in the literature (2, 4, 5, 7, 8). CEM indicated 43 of 44

TABLE 3 Sensitivity, specificity, accuracy, PPV and NPV levels according to BI-RADS assessment. 95% confidence intervals are presented in brackets.

	FFDM in CEM group	CEM	FFDM in DBT group	DBT
SENSITIVITY	40% [24%, 57%]	97% [88%, 99%]	55% [36%, 72%]	100% [89%, 100%]
SPECIFICITY	87% [83%, 93%]	63% [56%, 70%]	91% [85%, 95%]	60% [52%, 67%]
ACCURACY	80% [73%, 85%]	70% [64%, 76%]	85% [79%, 90%]	66% [59%, 72%]
PPV	44% [27%, 62%]	39% [30%, 49%]	53% [35%, 70%]	32% [23%, 42%]
NPV	87% [81%, 91%]	99% [95%, 99%]	91% [86%, 95%]	100% [97%, 100%]

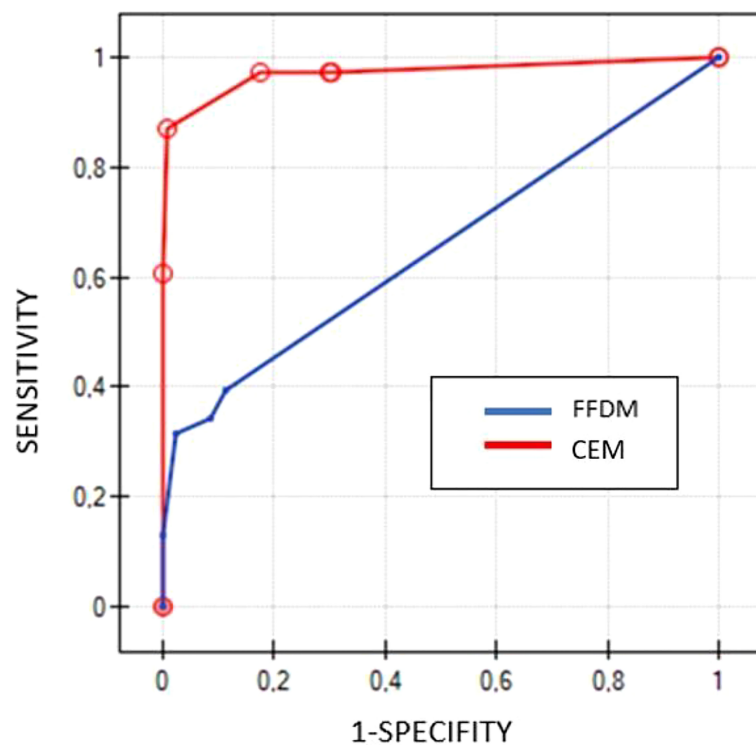


FIGURE 2

Comparison of ROC curves for CEM (red line) and FFDM (blue line) based on BIRADS scores.

histopathologically verified cancer cases in our group of patients. FFDM presented with a sensitivity of 40% and specificity of around 87%, indicating 15 of 44 cancer lesions. ROC curves presented significant differences ($p < 0.000001$) in diagnostic performance between CEM and FFDM (Figure 2). Despite the low concordance level of CEM and biopsy, it was slightly higher (0.39), as compared to FFDM and biopsy (0.29). The advantage of CEM over FFDM is obtained by the ability to visualize lesions obscured by dense glandular tissue. Moreover, the contrast enhancement allows to establish more accurately the size and number of lesions (1).

DBT allows viewing breast anatomy in multiple sections, which increases the diagnostic accuracy, as compared to FFDM, even if additional projections are employed (3, 9).

As predicted, the results of the DBT group also indicated an advantage of this technique over FFDM, with 100% sensitivity and specificity of 60%. DBT indicated all 33 cancers, confirmed in biopsy and there were no false-negative cases, whereas FFDM indicated 18 of them. The ROC curves, similarly to the CEM group, presented significant differences between DBT and FFDM ($p < 0.000001$) (Figure 3). Surprisingly, unlike in the CEM group, the kappa test indicated moderate concordance

between FFDM and biopsy (0.45), but low concordance between DBT and biopsy (0.32).

Despite the 100% accuracy of DBT and 97% accuracy of CEM, the kappa test indicated a low concordance level of both modalities with biopsy. Potential explanations include a high rate of false-positive cases in DBT (34%) and in CEM (30%). FFDM in the DBT group, despite a worse cancer detection rate, showed false results in 14% of cases, including 7% of false-negative and 7% of false-positive, which caused moderate concordance with biopsy (Figure 5).

It is interesting to note that the one lesion, which was indicated as a false-negative in CEM was invisible also in MR or PET/CT examinations. The patient was recalled due to enlarged axillary lymph nodes. Histopathological examination confirmed multicentric lobular cancer (EG: 75%, PG: 10%, HER2 negative, Ki67: 10%, NHG2) that might not be visible on imaging modalities such as CEM or MRI, due to its slow growth and low metabolic activity. (Figure 6) (5). It is not clear if the lesion would be visible in DBT. The patient had a multicentric invasive lobular cancer (ILC) which is characterized by slow growth and low metabolic activity. This is why the lesion was not detected by other modalities. DBT is

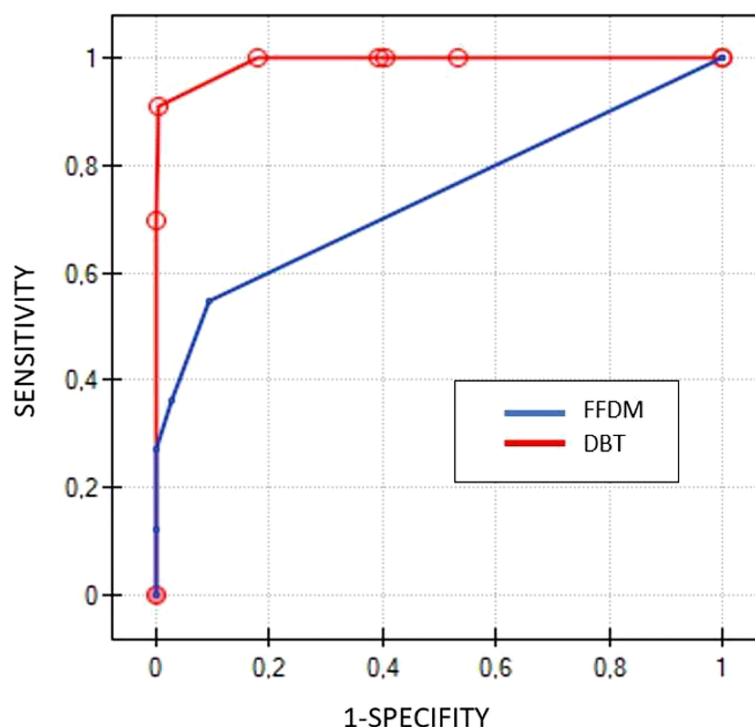


FIGURE 3

Comparison of ROC curves for DBT (red line) and FFDM (blue line) based on BIRADS scores.

the best imaging tool for the assessment of architectural distortions. Theoretically, large, slow growing ILC could manifest as a non-specific architectural distortion in DBT.

We found no significant differences between CEM and DBT (Figure 4). The groups were standardized, the patient's ages and sizes of lesions were similar. What is more, there were no significant differences in cancer detection rates. These results lead to the conclusion that CEM and DBT are equivalent and may be used alternatively in patients recalled from screening. One might expect that CEM would significantly outperform DBT, due to the administration of a contrast agent and subsequent independence of breast density (4, 5, 7, 8). Moreover, CEM provides data on tumor vasculature including neoangiogenesis (18, 19). In spite of these advantages, CEM cancer detection rate in our cohort did not outperform DBT. The probable explanation for that might be in the reduction of tissue overlap due to the tomographic technique of image acquisition (20). Furthermore, the radiation dose between CEM and DBT examinations in our study was similar.

On the contrary, authors of a recent paper who evaluated 220 women examined with CEM after being recalled from screening, found that CEM-based approach outperforms standard assessment (i.e. DBT or additional mammographic views) (21). However, the cited study was based on one-arm protocol and the age range of qualified patients was different

from ours, so the results are not exactly comparable. The age of 50-69 years is concordant with the European Commission Initiative on Breast Cancer guidelines of breast cancer screening. Also Girometti et al. presented a comparison of CEM and DBT+FFDM in 78 preoperative patients with histopathologically confirmed breast cancer and proved the superiority of CEM over DBT. However, it was a retrospective study comparing performance of 4 blinded readers, whereas our study is a prospective trial with a direct comparison of those two imaging tools (22). Moreover, Zuley et al. analyzed 60 lesions in 54 patients retrospectively and found that CEM reduced false-positive rate to 39% in comparison to FFDM/DBT (47%) and ultrasonography (61%) (23).

Due to the contrast injection, CEM examination is associated with a higher risk of complications, such as kidney failure or an allergic reaction. Approximately 2 in 10000 patients will develop anaphylaxis (0,15%) and 0,02-0,04% will develop severe anaphylaxis reaction. Contrast induced nephropathy develops in 1-2% patients. To prevent Post Contrast Acute Kidney Injury (PC-KAI) it is important to check eGFR level within three months before examinations and avoid repeated iodinated contrast examinations. The guidelines indicate that if the patient is under 70 years old, with no history of renal disease, proteinuria, hypertension, hyperuricemia or diabetes, then there is no need to check the eGFR level. During our study to prevent

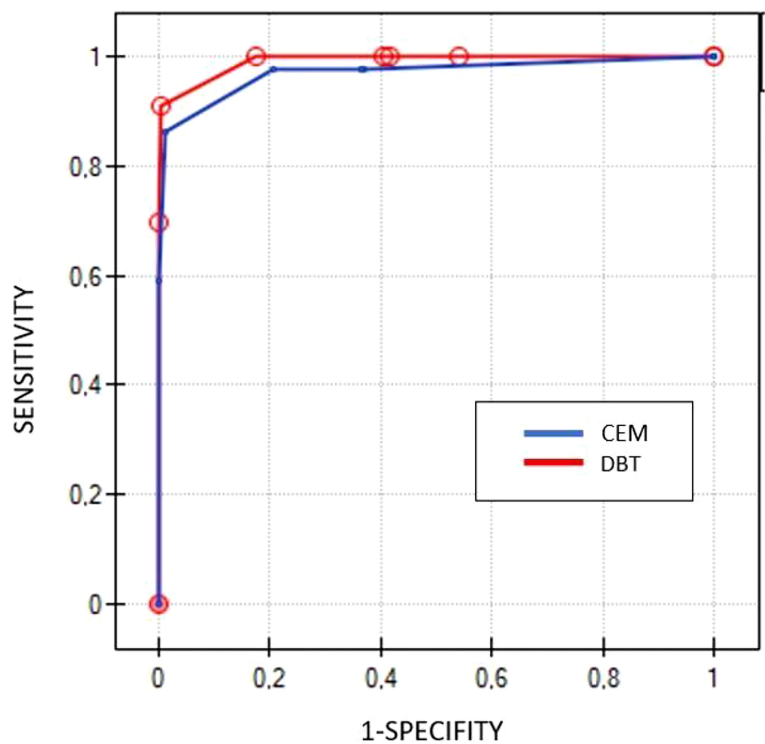


FIGURE 4
Comparison of ROC curves for DBT (red line) and CEM (blue line) based on BIRADS scores.

PC-AKI we evaluated creatinine and eGFR level in every patient (11, 24). On the other hand, DBT is more cost-effective and less stressful for patients (no contrast administration). Available literature focuses on the comparison of CEM or DBT with

FFDM independently. To the best of our knowledge, this is the first study comparing both modalities.

Our study has certain limitations. The major limitation is the low number of participating readers. There is a limited number

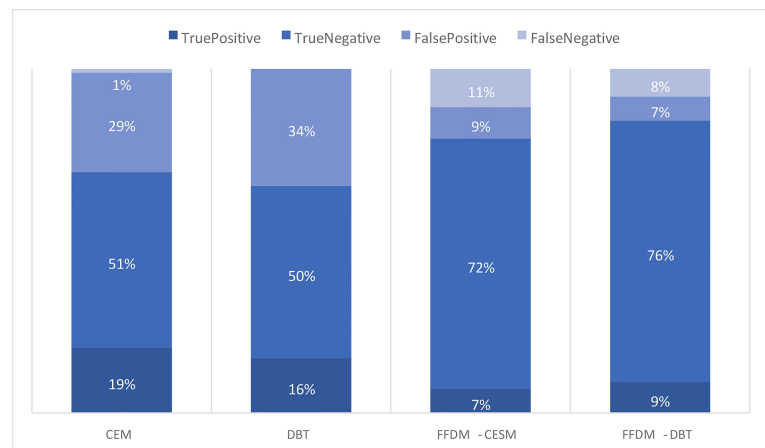


FIGURE 5
Comparison of the occurrence of true-positive, true-negative, false-positive and falsenegative results in CEM, DBT and FFDM subgroups.

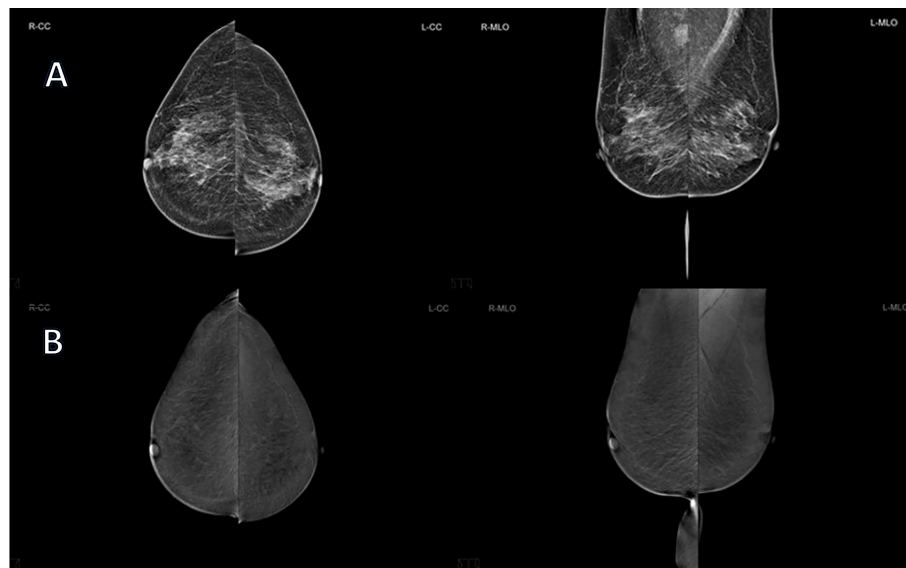


FIGURE 6

65 years old patient, with ACR type B breast density, recalled from screening due to the suspicious mass visible on RMLO in FFDM, classified as BIRADS 0 (A) and CEM (B) revealed no pathological enhancement. However, US was performed and it detected axillary lymph nodes with a typical metastatic appearance. One of them was biopsied and the pathology report revealed lymph node metastases of invasive lobular carcinoma. Breast MRI and PET-CT examinations were performed as well and they showed no signs of the primary tumor. The patient underwent neoadjuvant hormonal therapy and right mastectomy with axillary lymph node dissection. The final pathology report revealed multicentric invasive lobular carcinoma (grade 2) of the right breast with metastases to 23 out of 28 axillary lymph nodes.

of experienced breast radiologists in general, also in our center. This is a single-center study and therefore, despite the extensive experience of our radiologists participating in the evaluation of images, the results should be verified in a multicenter setting. Secondly, both techniques being compared, i.e. CEM and DBT, were used in parallel groups of patients. The reason for that trial design was related to the ethical considerations of avoiding additional risk (excessive radiation). Due to the risk of contrast-induced nephropathy, two patients with abnormal eGFR were excluded from the CEM group and DBT was performed instead.

Most of the studies on CEM in the diagnosis of breast cancer were performed using mammographs produced by GE Healthcare (8, 21). There is little data in literature on the performance of mammographs manufactured by Siemens that were used in our study. It may be attributed to the relatively recent appearance of Siemens CEM technology on the market. Another technological difference is the contrast agent used for CEM. Iopromide 370 (Ultravist 370) was used in our study, whereas other authors used Iohexol 300, Iohexol 350 or Iomeprol 400. All of these agents have slightly different concentrations of iodine (11).

In our opinion, larger multicenter trials should be taken into account to confirm our results in a larger, more heterogeneous population of patients recalled from screening. Such trials should be organized possibly on an international level in order

to establish the role of CEM and DBT in the management of this particular patient population independent of the influences of some local factors.

Conclusions

To our knowledge, this is the first report of a randomized comparison of CEM and DBT. Both studied modalities demonstrated high diagnostic performance, and none of them was found to be superior. Therefore, the choice of method should be based on availability, patient's safety and economic factors.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Regional Medical Bioethical Committee of Greater

Poland Medical Chamber. The patients/participants provided their written informed consent to participate in this study.

Author contributions

NS: Data analysis and interpretation. Writing the manuscript. AP-C: Data analysis and interpretation. Writing the manuscript. AG: Data collection and data analysis. PW: Data collection and statistical analysis. MO: Data collection. JM: Data collection. PN: Patients Management. PG: Patients Management. DG: Drafting the article. MR: Final approval. RC: Analysis and interpretation of results, drafting the article, final approval. All authors contributed to the article and approved the submitted version.

References

- Sardanelli F, Aase HS, Álvarez M, Azavedo E, Baarslag HJ, Balleyguier, et al. Position paper on screening for breast cancer by the European society of breast imaging (EUSOBI) and 30 national breast radiology bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, the Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey. *Eur Radiol* (2017) 27(7):2737–43. doi: 10.1007/s00330-016-4612-z
- FDA. 510(k) clearance for GE contrast enhanced spectral mammography (CESM). Available at: <https://www.accessdata.fda.gov/cdrhdocs/pdf10/K103485.pdf>.
- Dhamija E, Gulati M, Deo S, Gogia A, Hari S. Digital breast tomosynthesis: an overview. *Indian J Surg Oncol* (2021) 12(2):315–29. doi: 10.1007/s13193-021-01310-y
- Sogani J, Mango VL, Keating D, Sung JS, Jochelson MS. Contrast-enhanced mammography: Past, present, and future. *Clin Imaging* (2021) 69:269–79. doi: 10.1016/j.clinimag.2020.09.003
- James JJ, Tennant SL. Contrast-enhanced spectral mammography (CESM). *Clin Radiol* (2018) 73(8):715–23. doi: 10.1016/j.crad.2018.05.005
- Cozzi A, Magni V, Zanardo M, Schiaffino S, Sardanelli F. Contrast-enhanced mammography: A systematic review and meta-analysis of diagnostic performance. *Radiology* (2022) 302(3):568–81. doi: 10.1148/radiol.211412
- Zhu X, Huang JM, Zhang K, Xia LJ, Feng L, Yang P, et al. Diagnostic value of contrast-enhanced spectral mammography for screening breast cancer: systematic review and meta-analysis. *Clin Breast Cancer* (2018) 18:e985–95. doi: 10.1016/j.clbc.2018.06.003
- Łuczyńska E, Heinze-Paluchowska S, Hendrick E, Dyczek S, Ryś J, Herman K, et al. Comparison between breast MRI and contrast-enhanced spectral mammography. *Med Sci Monitor Int Med J Exp Clin Res* (2015) 21:1358–67. doi: 10.12659/MSM.893018
- Li T, Marinovich ML, Houssami N. Digital breast tomosynthesis (3D mammography) for breast cancer screening and for assessment of screen-recalled findings: review of the evidence. *Expert Rev Anticancer Ther* (2018) 18(8):785–91. doi: 10.1080/14737140.2018.1483243
- Zackrisson S, Lång K, Rosso A, Johnson K, Dustler M, Förnvik D, et al. One-view breast tomosynthesis versus two-view mammography in the malmö breast tomosynthesis screening trial (MBTST): A prospective, population-based, diagnostic accuracy study. *Lancet Oncol* (2018) 19(11):1493–503. doi: 10.1016/S1470-2045(18)30521-7
- Zanardo M, Cozzi A, Trimboli RM, Labaj O, Monti CB, Schiaffino S, et al. Technique, protocols and adverse reactions for contrast-enhanced spectral mammography (CESM): A systematic review. *Insights into Imaging* (2019) 10(1):1–15. doi: 10.1186/s13244-019-0756-0
- Dromain C, Balleyguier C, Muller S, Mathieu MC, Rochard F, Opolon P, et al. Evaluation of tumor angiogenesis of breast carcinoma using contrast-

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- enhanced digital mammography. *AJR Am J roentgenol* (2006) 187(5):W528–37. doi: 10.2214/AJR.05.1944
- D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. *ACR BI-RADS® atlas, breast imaging reporting and data system*. Reston, VA: American College of Radiology (2013).
- Lee CH, Phillips CJ, Sung JS, Lewin JM, Newell MS. *Contrast enhanced mammography 2022, a supplement to ACR BI-RADS® atlas, breast imaging reporting and data system*. Reston, VA: American College of Radiology (2013).
- NHSBSP Breast Screening Programme. *Guidelines for non-operative diagnostic procedures and reporting in breast cancer screening*. Sheffield: NHSBSP Publication No 50 (2001).
- Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. *Int J Cancer* (2021) 149(4):778–89. doi: 10.1002/ijc.33588
- Kornecki A. Current status of contrast enhanced mammography: A comprehensive review. *Can Assoc Radiologists J = J l'Association Can Des radiologistes* (2022) 73(1):141–56. doi: 10.1177/08465371211029047
- James JR, Pavlicek W, Hanson JA, Boltz TF, Patel BK. Breast radiation dose with CESM compared with 2D FFDM and 3D tomosynthesis mammography. *AJR Am J roentgenol* (2017) 208(2):362–72. doi: 10.2214/AJR.16.16743
- Jochelson MS, Lobbes M. Contrast-enhanced mammography: State of the art. *Radiology* (2021) 299(1):36–48. doi: 10.1148/radiol.2021201948
- Peppard HR, Nicholson BE, Rochman CM, Merchant JK, Mayo RC3rd, Harvey JA. Digital breast tomosynthesis in the diagnostic setting: Indications and clinical applications. *Radiographics* (2015) 35(4):975–90. doi: 10.1148/r.2015140204
- Cozzi A, Schiaffino S, Fanizza M, Magni V, Menicagli L, Monaco C, et al. Contrast-enhanced mammography for the assessment of screening recalls: a two-centre study. *Eur Radiol* (2022), 1–12. doi: 10.1007/s00330-022-08868-3
- Girometti R, Linda A, Conte P, Lorenzon M, De Serio I, Jerman K, et al. Multireader comparison of contrast-enhanced mammography versus the combination of digital mammography and digital breast tomosynthesis in the preoperative assessment of breast cancer. *La Radiologia Med* (2021) 126(11):1407–14. doi: 10.1007/s11547-021-01400-5
- Zuley ML, Bandos AI, Abrams GS, Ganott MA, Gizienski TA, Hakim C, et al. Contrast enhanced digital mammography (CEDM) helps to safely reduce benign breast biopsies for low to moderately suspicious soft tissue lesions. *Acad Radiol* (2020) 27(7):969–76. doi: 10.1016/j.acra.2019.07.020
- van der Molen AJ, Reimer P, Dekkers IA, Bongartz G, Bellin MF, Bertolotto M, et al. Post-contrast acute kidney injury—part 1: Definition, clinical features, incidence, role of contrast medium and risk factors. *Eur Radiol* (2018) 28(7):2845–55. doi: 10.1007/s00330-017-5246-5



OPEN ACCESS

EDITED BY

Laura Curiel,
University of Calgary, Canada

REVIEWED BY

Barbara Costantini,
Agostino Gemelli University Polyclinic
(IRCCS), Italy
Chukwuemeka Ihemelandu,
MedStar Washington Hospital Center,
United States

*CORRESPONDENCE

Eva K. Egger
eva-katharina.egger@ukbonn.de

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 02 October 2022

ACCEPTED 21 November 2022

PUBLISHED 15 December 2022

CITATION

Egger EK, Buchen MA, Recker F,
Stope MB, Strunk H, Mustea A and
Marinova M (2022) Predicting
incomplete cytoreduction in patients
with advanced ovarian cancer.
Front. Oncol. 12:1060006.
doi: 10.3389/fonc.2022.1060006

COPYRIGHT

© 2022 Egger, Buchen, Recker, Stope,
Strunk, Mustea and Marinova. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply
with these terms.

Predicting incomplete cytoreduction in patients with advanced ovarian cancer

Eva K. Egger^{1*}, Marie Antonia Buchen¹, Florian Recker¹,
Matthias B. Stope¹, Holger Strunk², Alexander Mustea¹
and Milka Marinova³

¹Department of Gynecology and Gynecological Oncology, University Hospital Bonn,
Bonn, Germany, ²Medicine Center Bonn, Medical Care Center, Bonn, Germany, ³Department of
Nuclear Medicine, University Hospital Bonn, Bonn, Germany

Purpose: The most important prognostic factor for survival in ovarian cancer patients is complete cytoreduction. The preoperative prediction of suboptimal cytoreduction, considered as any residual disease at the end of surgery, could prevent futile surgery and morbidity. Here, we aimed to identify markers in the preoperative abdominal CT scans of an unselected cohort of patients with ovarian cancer that are predictive of incomplete cytoreduction.

Methods: This is a single-institution retrospective analysis of 105 epithelial ovarian cancer (EOC) patients treated with surgical cytoreduction between 2010 and 2020. Twenty-two variables on preoperative abdominal CT scans were compared to the intraoperative macroscopic findings by Fisher's exact test. Parameters with a significant correlation between intraoperative findings and imaging were analyzed by multivariate binary logistic regression analysis regarding the surgical outcome of complete versus incomplete cytoreduction.

Results: Complete cytoreduction (CC), indicated by the absence of macroscopic residual disease, was achieved in 79 (75.2%) of 105 patients and 46 (63.9%) of 72 International Federation of Gynecology and Obstetrics (FIGO) stage III and IV patients. Twenty patients (19%) were incompletely cytoreduced due to miliary carcinomatosis of the small bowel, and six patients (5.7%) had various locations of residual disease. Thirteen variables showed a significant correlation between imaging and surgical findings. Large-volume ascites, absence of numerically increased small lymph nodes at the mesenteric root, and carcinomatosis of the transverse colon in FIGO stage III and IV patients decreased the rate of CC to 26.7% in the multivariate analysis.

Conclusion: Large-volume ascites, the absence of numerically increased small lymph nodes at the mesenteric root, and carcinomatosis of the transverse colon are markers in preoperative CT scans predicting a low chance for complete cytoreduction in unselected ovarian cancer patients in a real-world setting.

KEYWORDS

ovarian cancer, cytoreduction, CT scan, residual tumor, mesenteric lymph nodes

Introduction

Ovarian cancer is still the most frequent cause of death in women suffering from gynecologic malignancies (1). Standard treatment is upfront surgery followed by platinum- and taxane-containing chemotherapy. In the case of the International Federation of Gynecology and Obstetrics (FIGO) stage IIIB to IVB, bevacizumab, an antivascular endothelial growth factor antibody, is added. Additionally, patients with deficient homologous recombination are treated with PARP inhibitors (2–4). Optimal cytoreduction in epithelial ovarian cancer (EOC) patients, considered as no macroscopically visible residual disease at the end of surgery, is the most important factor for survival (2, 5–7). As most patients present in the advanced stages of the disease, optimal cytoreduction will include multivisceral surgery harboring the risk of morbidity and mortality (8). As reported rates of optimal cytoreduction range between 20% and 85%, there will be patients undergoing surgery without survival benefits and patients who might profit from neoadjuvant chemotherapy before debulking surgery (5). While the specificity of contrast-enhanced abdominal computed tomography (CT) scans for the detection of peritoneal carcinomatosis is about 88%, the sensitivity is only 68% (9). By far, the most common reason for suboptimal cytoreduction is extensive small bowel mesentery or serosal carcinomatosis, often underestimated in presurgical CT scans (10–12). Therefore, an optimal preoperative screening would identify the subgroup of patients where complete cytoreduction will not be possible in an upfront situation to avoid futile surgery.

In the case of recurrence, the prospectively validated Arbeitsgemeinschaft Gynäkologische Onkologie (AGO) score identifies 75% of patients with recurrent ovarian cancer where optimal cytoreduction will be achieved again. However, no such tool is available in the primary situation (13). In the primary setting, the therapeutic sequence—upfront surgery followed by adjuvant chemotherapy or neoadjuvant chemotherapy followed by surgery—remains the key issue arising in ovarian cancer patients deemed fit enough for surgery.

Here, we aimed to identify, in a real-life cohort of patients with ovarian cancer, the group of patients least likely to undergo complete cytoreduction despite a radical multivisceral surgical

approach by using radiological markers in the preoperative pelvic and abdominal CT scan.

Material and methods

Data collection

This study was conducted in accordance with the standards of the ethics committee of the Faculty of Medicine at the University of Bonn, Germany. The study obtained ethical approval (No 329/21) from the ethics committee of the Faculty of Medicine at the University of Bonn, Germany. Patients gave informed consent for the use of their data. The institutional record database was screened for epithelial ovarian cancer patients with cytoreductive surgery between January 2010 and December 2020. A total of 346 patients were identified. Patients with recurrent disease ($n = 63$) and patients with CT examinations without oral and intravenous contrast administration ($n = 178$) were excluded from the analysis. All CT scans were performed within a maximum of 28 days before surgery. In the case of neoadjuvant chemotherapy, two or three cycles of chemotherapy were completed before CT scan acquisition. Gastrografin was used as a radiopaque contrast medium 1 h prior to image acquisition; an intravenous contrast agent (iopamidol) was also administered. The CT scan was performed with the patients in a supine position by using a 64-slice scanner (Brilliance, Philips Healthcare, Amsterdam, the Netherlands); both arterial and portal venous phase images were acquired. Two radiologists with at least 15 years of experience in abdominal imaging blinded to the surgical details and outcome were asked to evaluate all abdominal CT scans for the following 23 items: liver metastasis, ascites, absence/presence of numerous small lymph nodes at the mesenteric root (number > 10 , short diameter < 1 cm), paracolic peritoneal carcinomatosis (PC), right and left diaphragm thickening as a sign of PC, general peritoneal thickening, PC of the small and large bowel mesentery, PC of the small bowel mesenteric root, PC of the spleen, extrahepatic PC considered as PC on Glisson's capsule, PC in the porta hepatis/hepatoduodenal ligament, PC of the gallbladder, wall thickening of the small bowel as suspected correlate of a serosal PC, PC of the

rectosigmoid, PC of the transverse colon, PC of the ileocecal region, PC at the omentum minus and majus, PC on the stomach wall, pelvic tumor, and retroperitoneal infrarenal lymph node enlargement.

Surgery reports and pathologic findings were screened for carcinomatosis in all the above-mentioned regions. Ascites were measured by CT scan only, as surgery reports were too vague. The peritoneal carcinomatosis index as the sum of carcinomatosis, quantified by size in 13 regions of the abdomen, was retrospectively calculated based on surgical and pathological reports to provide information on tumor burden (14). The main criterion for optimal debulking was no macroscopically visible residual disease at the end of surgery. In all cases of incomplete resection, the location of tumor residuals was documented.

Statistical analysis

In the first step, all variables were analyzed by Fisher's exact test to identify significant correlations between imaging and intraoperative finding. Differences were considered to be significant at a threshold of ≤ 0.05 . In a second step considering FIGO stage III and IV patients only, variables with a significant correlation of imaging and intraoperative finding were analyzed by multivariate binary logistic regression regarding the surgical outcome (complete or incomplete cytoreduction). The positive and negative predictive values of the CT scan were calculated for the analysis of serosal, mesenteric, and mesenteric root carcinomatosis. All statistical analyses were performed using Minitab Version 18 (Minitab LLC, State College, PA, USA).

Results

Baseline information

Baseline patient characteristics are shown in Table 1. Presurgical CT scans of 105 non-selected patients with ovarian cancer were evaluated. At surgery, complete cytoreduction (CC) was achieved in 79 (75.2%) of 105 patients and 46 (63.9%) of 72 FIGO stage III and IV patients. Twenty patients (19%) underwent incomplete cytoreduction due to miliary carcinomatosis of the small bowel; in further six patients (5.7%), the reason was tumor involvement of the porta hepatis ($n = 2$), liver metastases ($n = 1$), and spread to the retroperitoneum ($n = 2$) or pancreas ($n = 1$). Thirteen of 22 variables in total showed significant correlations between imaging and surgical findings as depicted in Table 2. The evaluation of the preoperative CT scans was especially difficult regarding the issue of carcinomatosis of the small and large bowel mesentery with 57 and 43 patients regarded as not evaluable. The 23rd variable "ascites", was only evaluated on

CT scans, as the surgical reports showed low accuracy regarding the three predefined conditions: 1) no ascites, 2) ascites only in the pelvis, and 3) ascites in all four quadrants of the abdomen. No ascites were seen in 51 patients, only in the pelvis in 16 patients, and ascites in all four quadrants of the abdomen were present in 38 patients.

Figures 1–3 are representative examples of evaluated CT findings within our real-world cohort of patients.

Small bowel carcinomatosis

As depicted in Table 2, there was a significant correlation between suspected small mesentery carcinomatosis in the preoperative CT scan and the intraoperative finding of a small bowel mesentery carcinomatosis ($p = 0.001$). However, 57 patients (54.3%) were deemed to be not evaluable in the preoperative CT scan regarding this issue. The negative predictive value (NPV) was 35%, and the positive predictive value (PPV) was only 7.14%. Surgery reports documented a small bowel mesentery carcinomatosis in 23 of 57 non-evaluable patients. The intraoperatively documented small bowel serosal carcinomatosis showed no significant correlation with the CT scan finding of a wall thickening of the small bowel ($p = 0.08$). The NPV was 87.67%, and the PPV was 25.59%. Small bowel carcinomatosis of the root was suspected in 48 cases, and in 55 cases, the root was deemed unsuspecting, and two cases were not evaluable. During surgery, 15 of 48 suspected patients showed carcinomatosis of the root of the small mesentery, and 5 of 56 unsuspecting patients showed a mesentery root carcinomatosis reaching a significant correlation ($p = 0.006$). The NPV was 90.91%, but the PPV was only 31.35%.

Looking into the visibility of lymph nodes within the mesenteric root, we found 51 patients in total with numerous (>10) small mesenteric lymph nodes at the mesenteric root (number > 10 , short axis diameter < 1 cm, oval configuration) and 53 patients without any visible mesenteric lymph node at the mesenteric root in the preoperative CT scan. One patient was deemed to be not evaluable using the preoperative CT scan.

The absence of multiple (>10) small mesenteric lymph nodes, defined as no visible mesenteric lymph nodes in the mesenteric root, as shown in Figure 4, was significantly more frequently observed in case of miliary carcinomatosis of the small bowel serosa, mesentery, or root, detected during surgery.

In detail, 40 of 67 patients without carcinomatosis of the mesentery showed numerous mesenteric root lymph nodes, while 27 of 38 patients with mesenteric carcinomatosis had no detectable mesenteric lymph node in the mesenteric root on the CT scan ($p = 0.002$), showing a PPV of 21.57% and an NPV of 50% for the detection of peritoneal carcinomatosis of the mesentery in the preoperative CT scan.

Of 21 patients with carcinomatosis of the mesenteric root, 18 showed no visible mesenteric lymph node, while 48 of 84

TABLE 1 General patient characteristics.

N	105
Age	Median 53 (range 32–82)
Histology	Serous: 91/86.7% Endometrioid: 9/8.6% Mucinous: 3/2.9% Clear cell: 2/1.9%
BMI	<19: 6/5.7% 20–24: 48/45.7% 25–30: 32/30.5% 31–40: 14/13.3% >40: 5/4.8%
Cytoreduction	Complete: 79/75.2% Incomplete: 26/24.8%
Neoadjuvant chemotherapy	Yes: 54/51.4% No: 51/48.6%
Duration of surgery in minutes	Median 343 (range 126–691)
Number of erythrocyte concentrates	Median 2 (range 0–19)
Ascites	None: 51/48.6% Only pelvic ascites: 16/15.2% In all 4 quadrants: 38/26.2%
Peritoneal carcinomatosis index (PCI)	Median 9 (range 1–29)
Preoperative CA 12-5	Median 115 (range 11–9,647 U/ml)
FIGO stage	No. of patients
IA	11
IB	1
IC	9
IIA	2
IIB	6
IIC	3
IIIA	5
IIIB	4
IIIC	55
IVA	6
IVB	2

BMI, body mass index; FIGO, International Federation of Gynecology and Obstetrics.

patients without carcinomatosis of the mesenteric root showed numerous mesenteric lymph nodes ($p < 0.001$), showing a PPV of 5.9% and an NPV of 66.7% for the detection of peritoneal carcinomatosis of the mesenteric root in the preoperative CT scan.

In 14 of 18 patients with small bowel mesenteric serosal carcinomatosis, detected during surgery, no mesenteric lymph nodes were detectable within the preoperative CT scan, while 47 of 87 patients without serosal carcinomatosis showed numerous mesenteric lymph nodes ($p = 0.019$), showing a PPV of 7.8% and an NPV of 74.1% for the detection of serosal peritoneal carcinomatosis in the preoperative CT scan.

In total, 40 patients had an intraoperatively documented carcinomatosis in at least one of the three above-mentioned locations (mesenteric root, small bowel mesentery, or small bowel serosa). In 28 of 40 patients, mesenteric lymph nodes were absent in the preoperative CT scan. In 12 patients, lymph

nodes were present within the mesenteric root in the preoperative CT scan ($p = 0.003$), showing a PPV of 52.8% and an NPV of 76.5% for the detection of peritoneal carcinomatosis of the small bowel at the mesenteric root, the mesentery, or the serosa in the preoperative CT scan. Figure 5 shows the intraoperative finding of carcinomatosis of the mesenteric root, the mesentery, and the serosa.

Large-volume ascites were found in 38 patients. Of these patients, 27 showed carcinomatosis of the small bowel in at least one of the above-mentioned locations (root, mesentery, or serosa). In 19 of 38 patients with large-volume ascites, complete cytoreduction was achieved.

Pathology reports in the case of large-volume ascites showed 34 high-grade serous, three high-grade endometrioid, and one clear cell histology.

An omental cake was seen in 33 patients. In 9 of 33 patients, the imaging showed no serosal or transmural transverse colon carcinomatosis, and in two cases, imaging was inconclusive regarding a serosal/transmural transverse colon infiltration. In further 10 patients, there was a transmural transverse colon infiltration by large tumor nodules originating from the colic mesentery in absence of an omental cake.

Cytoreduction was incomplete in 26 patients, and 21 of them showed miliary small bowel carcinomatosis (serosa, root, and/or mesentery), but in one patient, the irresectability was mainly due to involvement of the porta hepatis. In 17 of 21 patients, there were no small mesenteric lymph nodes visible. In total, small mesenteric lymph nodes were absent in 53 patients. In 28 of 53 patients, there was a small bowel carcinomatosis (serosa/root and/or mesentery) present.

Regression analysis of all parameters with significant correlation between intraoperative findings and preoperative CT scan

In FIGO stage III and IV patients, the completeness of cytoreduction was compared to 13 variables that achieved significance with respect to imaging and intraoperative finding and the presence of no, little, or large-volume ascites. In the case of ascites in all four quadrants of the abdomen, considered large-volume ascites with the absence of numerous small mesenteric lymph nodes and peritoneal carcinomatosis of the transverse colon, the rate of complete cytoreduction was as low as 26.68%. In the case of a tumor-free transverse colon in the preoperative CT scan but large-volume ascites and absent mesenteric lymph nodes, the rate for complete cytoreduction was as low as 44.51% as seen in Table 3.

The same group of patients was further distinguished into patients receiving neoadjuvant chemotherapy and interval debulking surgery (IDS) and patients receiving primary

TABLE 2 Correlation of tumor location according to CT scan and according to surgery report.

Tumor location	Imaging: tumor/ no tumor	Surgery reports: tumor/ no tumor	Imaging not conclusively assessable/no documented intraoperative assessment	p- Value
Paracolic	36/68	70/35	1/0	0.03
Right diaphragm	34/71	44/61	–	0.001
Left diaphragm	24/81	27/78	–	0.02
Pelvic tumor	87/18	82/23	–	0.004
Peritoneal thickening in general	78/25	30/75	2/0	0.002
Small bowel mesentery	20/28	38/67	57/0	0.001
Large bowel mesentery	20/42	38/67	43/0	<0.001
Mesenteric root	48/55	21/84	2/0	0.006
Splenic hilum	14/90	15/90	1/0	0.005
Liver surface	41/64	8/97	–	0.26
Porta hepatis	40/65	6/99	–	0.002
Gall bladder	10/78 (12× CHE)	2/91	5/0	>0.99
Infrarenal lymph nodes	104/1	30/74	0/1	>0.99
Small intestine serosa	29/73	18/87	3/1	0.08
Rectosigmoid	45/38	56/49	22/0	0.54
Transverse colon	32/69	22/83	4/0	0.002
Ileocecal pol	13/86	21/84	6/0	0.06
Stomach wall	18/86	8/97	1/0	0.003
Omental cake	33/71	52/53	1/0	<0.001
Liver metastases	4/100	4/100	1/0	0.15
Omentum minus	11/89	10/95	5/0	0.08

CHE, cholecystectomy.

debulking surgery (PDS). In the case of IDS large-volume ascites, absent numerous mesenteric lymph nodes and peritoneal carcinomatosis of the transverse colon led to a complete cytoreduction rate of 24.23% and in the case of PDS to a complete cytoreduction rate of 15.66% as depicted in Table 3.

FIGO stage I and II patients were excluded from the multivariate analysis, as complete cytoreduction is always possible in this patient population.

Discussion

In this study, the combination of large-volume ascites, peritoneal carcinomatosis of the transverse colon, and the absence of numerous small lymph nodes in the small bowel mesentery in the preoperative CT scans of the abdomen identified a group of patients where complete cytoreduction was achieved in less than 27% of the patients. Of course, optimal cytoreduction is a matter of tumor burden and surgical skill. When analyzing the optimal cytoreduction rate of ovarian cancer patients at our institution, we found that 75.2% of patients had no macroscopic visible residual tumor at the end of surgery. In the case of advanced disease (FIGO stage III and

IV), the rate was 63.9%, consistent with recently published data (2, 10, 11, 15). The most frequent site of failure of optimal cytoreduction in our study was a carcinomatosis of the small bowel (mesentery, root, and serosa) accounting for almost 80% of all cases with residual tumors. Similar findings were previously described, concluding that the success of surgery regarding optimal cytoreduction in ovarian cancer patients depends on the presence or absence of PC in specific regions rather than only on the amount of PC in general (10, 11, 16).

While preoperative staging by computed tomography of the abdomen and thorax is by far the most common approach for presurgical evaluation due to its wide availability, the substantial underestimation of visceral small bowel peritoneal carcinomatosis is its major drawback (16, 17). The general pooled sensitivity and specificity for correct identification of region-based peritoneal carcinomatosis is 68% and 88% in ovarian and gastric cancers, respectively, depending on the size of the lesions and the presence of ascites (9).

As the intraoperatively generated PCI score shows low interobserver variability, several promising attempts were made to describe tumor load by a preoperative CT scan-based PCI score. It was limited by a general underestimation of the tumor burden in the small bowel and hepatoduodenal ligament

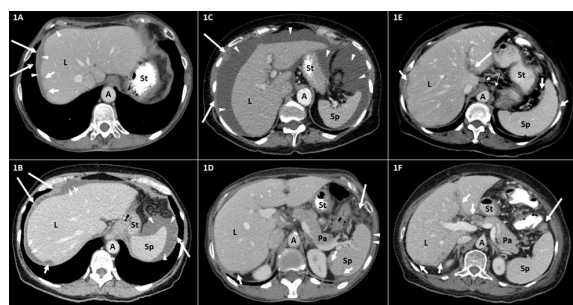


FIGURE 1

Representative CT findings in the upper abdomen in patients with advanced ovarian carcinoma FIGO stage III–IV. Axial contrast-enhanced CT scans of the upper abdomen (A–F). A, aorta; L, liver; Pa, pancreas; Sp, spleen; St, stomach. (A, B) Tumor implants of the diaphragm (long arrows) and the liver (short arrows). The thickening of the right hemidiaphragm (long arrows) can be distinguished from surrounding perihepatic ascites (arrowheads). The surface tumor deposit at the dome of the liver (short arrows) causes scalloping of the lateral (A) and posterior (B) liver surface. (C) Large amount of ascites (arrowheads) in the upper abdomen. Peritoneal knotty implants (long arrows) are shown. (D) Tumor scalloping (short arrows) of the posterior surface of the spleen and the liver. Small amount of perisplenic ascites (arrowheads). Fat tissue stranding and tumor nodularity (long arrow) are seen in the fat adjacent to the splenic flexure of the colon. (E) Tumor implants (long arrow) along the hepatogastric ligament. Peritoneal thickening and tumor scalloping of the surface of the spleen and the liver (short arrows). Round suprarenal lymph nodes (narrow arrows). (F) Tumor implants (short arrows) in the porta hepatis along the falciform ligament and of the posterior liver surface. Tumor nodule (long arrow) is seen in the fat adjacent to the splenic flexure of the colon. FIGO, International Federation of Gynecology and Obstetrics.

(regions 2 and 9 to 12), which are the most likely locations for residual disease (10, 11, 14, 18–20).

The CT scan sensitivity decreases substantially in tumor sizes below 5 mm. In our experience, the size of 5 mm or less comprises the size of the single carcinomatosis nodule on the visceral peritoneum seen during surgery in most cases (16, 21). Therefore, the small bowel carcinomatosis itself, the most common location of residual disease, is usually not visible in the presurgical CT scan.

In this context, MRI (contrast-enhanced and diffusion-weighted imaging (DWI) images) is generally considered more accurate and sensitive, especially for the detection of liver metastases, perihepatic and serosal tumor nodules, and tumor implants on the hepatoduodenal and gastrohepatic ligament, diaphragm, and small intestinal wall. In contrast, results of recent studies showed that despite the highest sensitivity of MRI and the highest specificity of FDG-PET/CT, no significant differences were found between the three techniques (MRI, CT, and FDG-PET/CT) (22). Therefore, as the fastest, most economical, and widely available modality in daily practice and real life, CT is the examination of choice in particular when a

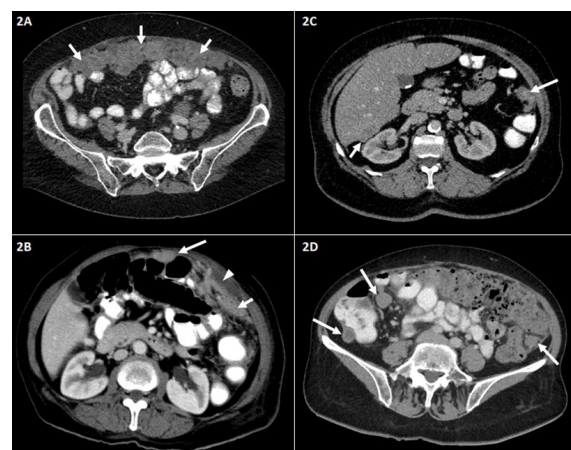


FIGURE 2

Representative abdominal CT findings in patients with advanced ovarian carcinoma FIGO stage III–IV. Axial contrast-enhanced CT scans of the middle and lower abdomen (A–D). (A) Irregular soft tissue mass representing large omental plaques called omental cake (short arrows) common site of intraperitoneal seeding of ovarian carcinoma. (B) Peritoneal thickening (short arrow) and small amount of ascites (arrowhead). Peritoneal nodule adjacent to the transverse colon (long arrow). (C) Peritoneal nodules in the left paracolic gutter (long arrow) and the hepatorenal recess (Morison's pouch) (short arrow). (D) Peritoneal implants/nodules in the right and left paracolic gutter. FIGO, International Federation of Gynecology and Obstetrics.

stand-alone technique is needed. If inconclusive, PET/CT or MRI may offer additional insights. Whole-body FDG-PET/CT may be more accurate for a supradiaphragmatic metastatic extension. Despite advances in imaging techniques, neither DWI-MRI nor CT nor FDG-PET/CT seems to be superior in preoperative assessment of the surgical PCI in patients scheduled for upfront cytoreductive surgery for advanced-stage EOC patients (23).

As there are numerous patients with high tumor load but tumor-free small bowel and tumor-free hepatoduodenal ligaments, patients with the residual disease may represent a subgroup of epithelial ovarian cancer patients with increased tumorigenicity, which allows this unfavorable unresectable tumor spread pattern (24). In our experience, carcinomatosis of the small bowel was always accompanied by a high tumor load (PCI > 15). The absence of multiple mesenteric lymph nodes in the case of peritoneal carcinomatosis of the small bowel may be due to a decreased immune reaction. In this context, it is interesting to note that in triple-negative breast cancer patients, a missing germinal center formation in cancer-free lymph nodes is an indicator of a poorer prognosis. Therefore, this was considered a sign of a decreased systemic immune response (25). Furthermore, MRI in the case of Crohn's disease remains unspecific regarding the proximal disease extension. However, the evaluation of the inflammation of small bowel mesentery lymph nodes shows the proximal disease extent despite

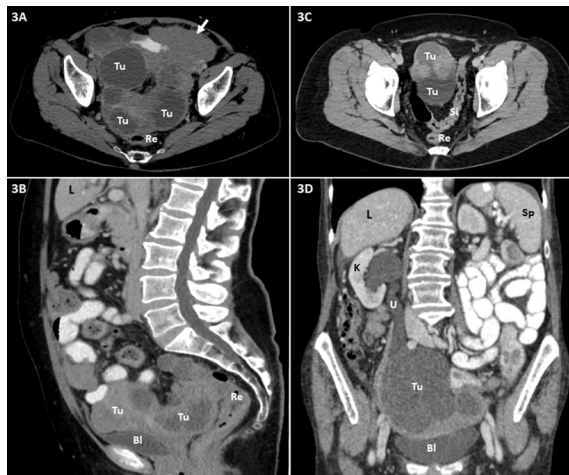


FIGURE 3

Abdominopelvic contrast-enhanced CT scans (A and C, axial images; B, sagittal reformatted image; D, coronal reformatted image) of patients with advanced ovarian carcinoma are shown. Bl, urinary bladder; L, liver; K, kidney; Re, rectum; Si, sigmoid colon; Sp, spleen; Tu, tumor (ovarian carcinoma); U, ureter. (A, B) Extensive tumor (Tu) in the pelvis consisting of large cystic and partly solid nodular parts. Close positional relationship and contact of the mass to the dorsally adjacent rectum (Re) and ventrally adjacent urinary bladder (Bl) with suspected rectal invasion and peritoneal involvement at the urinary bladder roof. Omental caking (short arrow) in the left side of the upper abdomen. (C) Massive tumor (Tu) in the pelvis including cystic and solid nodular parts. Close positional relationship and contact of the mass to the adjacent sigmoid colon (Si); tumor involvement of the latter is conceivable. (D) Large, predominantly cystic tumor mass (Tu) in the pelvis with involvement of the right ureter (U) and consecutive ipsilateral urinary retention and hydronephrosis of the right kidney (K).

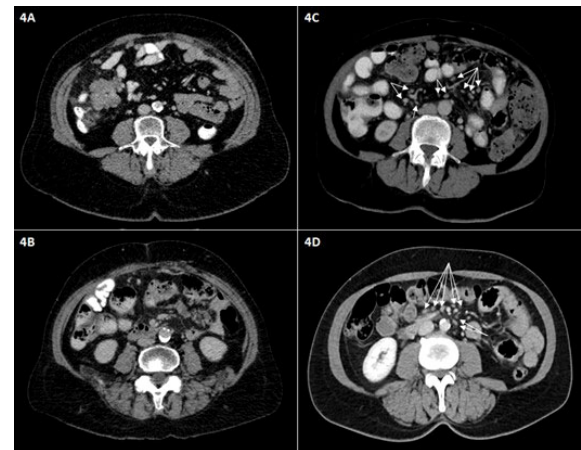


FIGURE 4

Representative axial intravenous contrast-enhanced CT scans of the middle abdomen. The absence of numerous small mesenteric lymph nodes is shown (A, B); only small vessels in the mesenteric root are visible. (C, D) Other patients show multiple small mesenteric lymph nodes (<1 cm in short dimension) in the mesentery and mesenteric root (arrows).

unsuspicious bowel walls. Considering the visible inflammation in the case of an active Crohn's disease, invisible mesenteric lymph nodes in the case of miliary small bowel carcinomatosis seem a noteworthy feature and a possible example of an immune escape of the tumor, leading to this unfavorable tumor spread (26). In addressing the problem of a diffuse tumor spread on the guts, three different clinical phenotypes of epithelial ovarian cancer patients were recently introduced, defining the diffuse tumor spread pattern within the rectosigmoid mesentery as the most lethal phenotype as compared to two other phenotypes with more localized disease and better survival outcomes (27).

Several studies reported diffuse peritoneal thickening, mesenteric disease, suprarenal lymph nodes, large-volume ascites, and carcinomatosis on the diaphragm or liver as significant markers in their final prediction model for complete cytoreduction (5, 21, 28, 29). The implementation of extensive upper abdominal surgery including diaphragm stripping, splenectomy, distal pancreatectomy, and resection of disease on the hepatoduodenal ligament made some of the above-mentioned markers less predictive for complete cytoreduction (10, 11, 15, 21, 28, 29).



FIGURE 5

Intraoperative finding of peritoneal carcinomatosis of the mesenteric root.

TABLE 3 Multivariate binary logistic regression analysis of 14 significant variables in FIGO IIIA–IVB patients.

Ascites 1. .none 2. .only pelvic ascites 3. . .ascites in all 4 quadrants	Numerous small mesenteric lymph nodes 0. .invisible 1. .visible	Involvement of the transverse colon 1. .yes 2. .no	Rate of completely cytoreduced patients in total	Rate of completely cytoreduced patients in case of primary debulking surgery (PDS)	Rate of completely cytoreduced patients in case of interval debulking surgery (IDS)
1	0	1	68.82%	29.78%	68.23%
1	0	2	82.94%	64.95%	79.79%
1	1	1	89.5%	54.90%	88.37%
1	1	2	94.94%	84.14%	93.39%
2	0	1	56.15%	22.23%	55.79%
2	0	2	73.84%	55.53%	70.11%
2	1	1	83.19%	45.07%	81.70%
2	1	2	91.60%	78.19%	89.24%
3	0	1	26.68%	15.56%	24.23%
3	0	2	44.51%	44.79%	37.40%
3	1	1	58.44%	34.77%	53.11%
3	1	2	75.60%	69.96%	67.88%

Laparoscopy is an interesting evaluation tool in advanced ovarian cancer patients with positive prediction rates of 62% for cytoreduction to less than 1 cm of tumor rest (10, 30–33). To optimize that rate, different calculation models included the findings of CT scans, laparoscopy, and laparotomy, identifying a marker constellation of a PCI of 20 and more and bowel obstruction as significant for incomplete cytoreduction. As bowel obstruction is rather rare despite high tumor loads and reason for surgical intervention anyhow, this model may not be suitable for all patients in everyday business to prevent futile surgery (34).

Of course, there are limitations to our study. First of all, its retrospective nature always harbors the risk of bias. Second, there is a relatively high rate of patients treated with neoadjuvant chemotherapy due to the prior policy of the clinic to administer neoadjuvant chemotherapy in case of ascites of more than 500 cc.

The strength of our study compared to others is the definition of macroscopically no residual disease as an optimal cytoreductive outcome as residual disease regardless of whether size impacts prognosis more severely than any further available therapeutic tool (2, 12). Furthermore, we analyzed only radiological features, as we believe that incomplete cytoreduction due to a compromised performance status or ASA status as well as age comprises a different group of patients generally undergoing incomplete resection to avoid complications rather than because of an unfavorable tumor spread pattern (35). A third noteworthy feature is that we included only primary ovarian cancer patients, as prior surgery will affect the quality of the CT scan evaluation (16). A final strength of our study is that no further surgical intervention, anesthesia, or a new method of imaging was

necessary to identify this unfavorable group of patients, which makes our finding applicable in the clinical day-to-day business (36).

To our knowledge, we are the first to describe the absence of multiple small bowel mesenteric lymph nodes as one of three markers predicting a very low chance for complete cytoreduction. As only one diagnostic tool was necessary for our finding, it is noteworthy for further prospective evaluation in larger cohorts and in combination with additional diagnostic techniques to encircle the group of patients without or with little benefit from surgery even better. Therefore, our finding may add valuable information for the decision between neoadjuvant chemotherapy or upfront debulking in ovarian cancer patients deemed fit enough for surgery. In our department, we implemented laparoscopy as an additional diagnostic tool in case of large-volume ascites, carcinomatosis of the transverse colon, and absent small bowel mesenteric lymph nodes in the preoperative CT scan. Different algorithms as additional preoperative abdominal MRI seem worth studying as well, in order to minimize any surgical intervention.

The key finding of our analysis is that the absence of numerous small mesenteric lymph nodes in the context of large-volume ascites and peritoneal carcinomatosis of the transverse colon in the presurgical CT scan of ovarian cancer patients are highly suspicious for miliary carcinomatosis of the small bowel and an unresectable tumor spread pattern.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics committee of the Faculty of Medicine at the University of Bonn, Germany. The patients/participants provided their written informed consent to participate in this study.

Author contributions

EE and MM conceived and designed the analysis. MB, MM, and HS collected the data. FR and MS contributed data or analysis tools. MB and EE performed the analysis. EE wrote the paper. AM and MM edited the paper. All authors contributed to the article and approved the submitted version

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
2. Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: A combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft gynaekologische onkologie studien-gruppe ovarialkarzinom (AGO-OVAR) and the groupe d'Investigateurs nationaux pour les études des cancers de l'Ovaire (GINECO). *Cancer* (2009) 115:1234–44. doi: 10.1002/cncr.24149
3. Ray-Coquard I, Pautier P, Pignata S, Pérol D, González-Martín A, Berger R, et al. PAOLA-1 investigators. olaparib plus bevacizumab as first-line maintenance in ovarian cancer. *N Engl J Med* (2019) 381:2416–28. doi: 10.1056/NEJM0A1911361
4. Heitz F, Harter P, Barinoff J, Beutl B, Kannisto P, Grabowski JP, et al. Bevacizumab in the treatment of ovarian cancer. *Adv Ther* (2012) 29:723–35. doi: 10.1007/s12325-012-0041-9
5. Bristow RE, Tomacruz RS, Armstrong DK, Trimble EL, Montz FJ. Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: A meta-analysis. *J Clin Oncol Am Soc Clin Oncol (ASCO)* (2002) 20:1248–59. doi: 10.1200/JCO.20.5.1248
6. Chang SJ, Hodeib M, Chang J, Bristow RE. Survival impact of complete cytoreduction to no gross residual disease for advanced-stage ovarian cancer: A meta-analysis. *Gynecol Oncol* (2013) 130:493–8. doi: 10.1016/j.ygyno.2013.05.040
7. Eisenhauer EL, Abu-Rustum NR, Sonoda Y, Aghajanian C, Barakat RR, Chi DS. The effect of maximal surgical cytoreduction on sensitivity to platinum-taxane chemotherapy and subsequent survival in patients with advanced ovarian cancer. *Gynecol Oncol* (2008) 108:276–81. doi: 10.1016/j.ygyno.2007.10.022
8. Egger EK, Kohls N, Stope MB, Condit M, Keyver-Paik MD, Könsgen D, et al. Risk factors for severe complications in ovarian cancer surgery. *In Vivo* (2020) 34:3361–5. doi: 10.21873/INVIVO.12174
9. van 't Sant I, Engbersen MP, Bhairosing PA, Lambregts DMJ, Beets-Tan RGH, van Driel WJ, et al. Diagnostic performance of imaging for the detection of peritoneal metastases: A meta-analysis. *Eur Radiol* (2020) 30(6):3101–12. doi: 10.1007/S00330-019-06524-X
10. Heitz F, Harter P, Alesina PF, Walz MK, Lorenz D, Groeben H, et al. Pattern of and reason for postoperative residual disease in patients with advanced ovarian cancer following upfront radical debulking surgery. *Gynecol Oncol* (2016) 141:264–70. doi: 10.1016/J.YGYNO.2016.03.015
11. Rosendahl M, Harter P, Bjørn SF, Høgdall C. Specific regions, rather than the entire peritoneal carcinosis index, are predictive of complete resection and survival in advanced epithelial ovarian cancer. *Int J Gynecol Cancer* (2018) 28:316–22. doi: 10.1097/IGC.0000000000001173
12. Harter P, Petran D, Scambia G, Ortega E, Tsubulak I, Nagao S, et al. Efficacy of maintenance olaparib plus bevacizumab by biomarker status in clinical higher-

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- and lower-risk patients with newly diagnosed, advanced ovarian cancer in the PAOLA-1 trial. *Int J Gynecol Cancer* (2020) 30:Suppl 3:A1–A210. doi: 10.1136/ijgc-2020-IGCS.18
13. Du Bois A, Sehouli J, Vergote I, Ferron G, Reuss A, Meier W, et al. Randomized phase III study to evaluate the impact of secondary cytoreductive surgery in recurrent ovarian cancer: Final analysis of AGO DESKTOP III/ENGOT-ov20. *JCO* (2020) 38(15-suppl):6000. doi: 10.1200/JCO.2020.38.15_SUPPL.6000
 14. Jacquet P, Sugarbaker PH. Clinical research methodologies in diagnosis and staging of patients with peritoneal carcinomatosis. *Cancer Treat Res* (1996) 82:359–74. doi: 10.1007/978-1-4613-1247-5_23
 15. Eisenhauer EL, Abu-Rustum NR, Sonoda Y, Levine DA, Poyner EA, Aghajanian C, et al. The addition of extensive upper abdominal surgery to achieve optimal cytoreduction improves survival in patients with stages IIIC-IV epithelial ovarian cancer. *Gynecol Oncol* (2006) 103:1083–90. doi: 10.1016/J.YGYNO.2006.06.028
 16. de Bree E, Koops W, Kröger R, van Ruth S, Witkamp AJ. Zoetmulder peritoneal carcinomatosis from colorectal or appendiceal origin: correlation of preoperative CT with intraoperative findings and evaluation of interobserver agreement. *J Surg Oncol* (2004) 86(2):64–73. doi: 10.1002/JSO.20049
 17. Michielsen K, Vergote I, de Bree KO, Amant F, Leunen K, Moerman P, et al. Whole-body MRI with diffusion-weighted sequence for staging of patients with suspected ovarian cancer: a clinical feasibility study in comparison to CT and FDG-PET/CT. *Eur Radiol* (2014) 24:889–901. doi: 10.1007/S00330-013-3083-8
 18. Gouy S, Belghiti J, Uzan C, Canlorbe G, Gauthier T, Morice P. Accuracy and reproducibility of the peritoneal cancer index in advanced ovarian cancer during laparoscopy and laparotomy. *Int J Gynecol Cancer* (2013) 23:1699–703. doi: 10.1097/IGC.0B013E3182A616A7
 19. Elias D, Souadka A, Fayard F, Mauguen A, Dumont F, Honore C, et al. Variation in the peritoneal cancer index scores between surgeons and according to when they are determined (before or after cytoreductive surgery). *Eur J Surg Oncol* (2012) 38:503–8. doi: 10.1016/J.EJSO.2012.01.001
 20. Koh JL, Yan TD, Glenn D, Morris DL. Evaluation of preoperative computed tomography in estimating peritoneal cancer index in colorectal peritoneal carcinomatosis. *Ann Surg Oncol* (2009) 16:327–33. doi: 10.1245/S10434-008-0234-2
 21. Dowdy SC, Mullany SA, Brandt KR, Huppert BJ, Cliby W. The utility of computed tomography scans in predicting suboptimal cytoreductive surgery in women with advanced ovarian carcinoma. *Cancer* (2004) 101:346–52. doi: 10.1002/CNCR.20376
 22. Schmidt S, Meuli RA, Achdari C, Prior JO. Peritoneal carcinomatosis in primary ovarian cancer staging: comparison between MDCT, MRI, and 18F-FDG PET/CT. *Clin Nucl Med* (2015) 40:371–7. doi: 10.1097/RLU.0000000000000768
 23. Mikkelsen MS, Petersen LK, Blaakaer J, Marinovskij E, Rodenkilde M, Andersen G, et al. Assessment of peritoneal metastases with DW-MRI, CT, and FDG PET/CT before cytoreductive surgery for advanced stage epithelial ovarian cancer. *Eur J Surg Oncol* (2021) 47:2134–41. doi: 10.1016/J.EJSO.2021.03.239

24. Drakes ML, Stiff PJ. Ovarian cancer: Molecular & diagnostic imaging and treatment strategies. *Adv Exp Med Biol* (2021) 1330:33–54. doi: 10.1007/978-3-030-73359-9_3
25. Liu F, Hardiman T, Wu K, Quist J, Gazinska P, Ng T, et al. ARTICLE systemic immune reaction in axillary lymph nodes adds to tumor-infiltrating lymphocytes in triple-negative breast cancer prognostication. *Breast Cancer* (2021) 7:1–10. doi: 10.1038/s41523-021-00292-y
26. Maconi G, Sorin V, Kopylov U, Barzilay O, Ferretti F, Innamorati S, et al. Diagnostic significance of mesenteric lymph node involvement in proximal small bowel crohn's disease. *Therap Adv Gastroenterol* (2022) 15:1–8. doi: 10.1177/17562848221118664
27. Park SJ, Mun J, Lee EJ, Sunwoo P, Sang YK, Whasun L, et al. Clinical phenotypes of tumors invading the rectosigmoid colon affecting the extent of debulking surgery and survival in advanced ovarian cancer. *Front Oncol* (2021) 11:673631. doi: 10.3389/FONC.2021.673631
28. Rutten MJ, van de Vrie R, Bruining A, Spijkerboer AM, Mol BW, Kenter GG, et al. Predicting surgical outcome in patients with international federation of gynecology and obstetrics stage III or IV ovarian cancer using computed tomography: A systematic review of prediction models. *Int J Gynecol Cancer* (2015) 25:407–15. doi: 10.1097/IGC.0000000000000368
29. Axtell AE, Lee MH, Bristow RE, Dowdy SC, Cliby WA, Raman, et al. Multi-institutional reciprocal validation study of computed tomography predictors of suboptimal primary cytoreduction in patients with advanced ovarian cancer. *J Clin Oncol* (2007) 25:384–9. doi: 10.1200/JCO.2006.07.7800
30. Fagotti A, Ferrandina G, Fanfani F, Garganese G, Vizzielli G, Carone V, et al. Prospective validation of a laparoscopic predictive model for optimal cytoreduction in advanced ovarian carcinoma. *Am J Obstet Gynecol* (2008) 199:642.e1–6. doi: 10.1016/j.AJOG.2008.06.052
31. Fagotti A, Vizzielli G, Fanfani F, Costantini B, Ferrandina G, Gallotta V, et al. Introduction of staging laparoscopy in the management of advanced epithelial ovarian, tubal and peritoneal cancer: impact on prognosis in a single institution experience. *Gynecol Oncol* (2013) 131:341–6. doi: 10.1016/j.YGYNO.2013.08.005
32. Fagotti A A, Ferrandina G, Fanfani F, Ercoli A, Lorusso D, Rossi M, et al. A laparoscopy-based score to predict surgical outcome in patients with advanced ovarian carcinoma: A pilot study. . *Ann Surg Oncol* (2006) 13:1156–61. doi: 10.1245/ASO.2006.08.021
33. Fleming ND, Westin SN, Meyer LA, Shafer A, Rauh-Hain JA, Onstad M, et al. Correlation of surgeon radiology assessment with laparoscopic disease site scoring in patients with advanced ovarian cancer. *Int J Gynecol Cancer* (2021) 31:92–7. doi: 10.1136/IJGC-2020-001718
34. Lluca A, Climent MT, Escrig J, Carrasco P, Serra A. MUAPOS working group (Multidisciplinary unit of abdominal pelvic oncology surgery) validation of three predictive models for suboptimal cytoreductive surgery in advanced ovarian cancer. *Sci Rep* (2021) 11:8111. doi: 10.1038/S41598-021-86928-2
35. Dessapt AL, Huchon C, Ngo C, Bats AS, Bensaid C, Lecuru F. Is complete cytoreductive surgery feasible in this patient with ovarian cancer? *Surg Oncol* (2016) 25:326–31. doi: 10.1016/J.SURONC.2016.07.001
36. Ghisoni E, Katsaros D, Maggiorotto F, Aglietta M, VairaM, De Simone M, et al. A predictive score for optimal cytoreduction at interval debulking surgery in epithelial ovarian cancer: A two-centers experience. *J Ovarian Res* (2018) 11:42. doi: 10.1186/s13048-018-0415-y



OPEN ACCESS

EDITED BY
Laura Curiel,
University of Calgary, Canada

REVIEWED BY
Firas Hamdan,
University of Maryland, Baltimore
United States
Deniz Can Guven,
Hacettepe University, Turkey

*CORRESPONDENCE
Mitsuaki Tatsumi
✉ m-tatsumi@radiol.med.osaka-u.ac.jp

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 24 August 2022
ACCEPTED 07 December 2022
PUBLISHED 22 December 2022

CITATION
Tatsumi M, Soeda F, Naka S,
Kurimoto K, Ooe K, Fukui H,
Katayama D, Watabe T, Kato H and
Tomiyama N (2022) Advantages of
FBPA PET in evaluating early response
of anti-PD-1 immunotherapy in
B16F10 melanoma-bearing mice:
Comparison to FDG PET.
Front. Oncol. 12:1026608.
doi: 10.3389/fonc.2022.1026608

COPYRIGHT
© 2022 Tatsumi, Soeda, Naka,
Kurimoto, Ooe, Fukui, Katayama,
Watabe, Kato and Tomiyama. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Advantages of FBPA PET in evaluating early response of anti-PD-1 immunotherapy in B16F10 melanoma-bearing mice: Comparison to FDG PET

Mitsuaki Tatsumi^{1*}, Fumihiko Soeda², Sadahiro Naka²,
Kenta Kurimoto², Kazuhiro Ooe², Hideyuki Fukui³,
Daisuke Katayama², Tadashi Watabe², Hiroki Kato²
and Noriyuki Tomiyama^{1,3}

¹Department of Radiology, Osaka University Hospital, Suita, Osaka, Japan, ²Department of Nuclear Medicine and Tracer Kinetics, Osaka University Graduate School of Medicine, Suita, Osaka, Japan, ³Department of Radiology, Osaka University Graduate School of Medicine, Suita, Osaka, Japan

Purpose: PET with L-4-borono-2-[¹⁸F] fluoro-phenylalanine (FBPA) was reported to be useful to differentiate malignant tumors and inflammation. Although immunotherapy with immune checkpoint inhibitors (ICIs) has been applied to cancer treatment recently, FDG PET may not be suitable to determine the effect of ICIs because of false-positive findings caused by treatment-related inflammation. In this study, we aimed to demonstrate that FBPA PET allowed detection of the early response of anti-PD-1 immunotherapy in tumor-bearing mice, comparing the results with those of FDG PET.

Materials and methods: Mice with B16F10 melanoma tumor xenografts were prepared. Anti-mouse PD-1 antibody or PBS was administered twice intraperitoneally to the tumor-bearing mice on Day 0 (3 days after inoculation) and Day 5 (treatment or control group <TrG or CoG>). PET/CT imaging was performed twice for each mouse on Day 0 before the anti-PD-1 antibody/PBS administration and on Day 7 using a micro-PET/CT scanner. FBPA and FDG PET/CT studies were conducted separately. SUVmax and the tumor to liver ratio (T/L ratio) were used as parameters exhibiting tumor activity. Tumor uptake volume (TUV) and metabolic tumor volume (MTV) were also calculated for FBPA and FDG, respectively. Changes between pre- and posttreatment SUVmax or T/L ratio were observed using the formula as follows: [(posttreatment parameter values/prettreatment values - 1) × 100] (%).

Results: Tumors in TrG were smaller than those in CoG on Day 7. SUVmax and T/L ratio represented no differences between TrG and CoG in FBPA and FDG PET before treatment. FBPA PET on Day 7 demonstrated that SUVmax, T/L ratio, and TUV in TrG were statistically smaller than those in CoG. %T/L ratio and %SUVmax exhibited the same trend in FBPA PET. However, FDG PET on Day 7

revealed no differences in all parameters between TrG and CoG. T/L ratio and % SUVmax in TrG represented larger values than those in CoG without statistical significances.

Conclusion: This study demonstrated that FBPA PET allowed detection of the early response of anti-PD-1 immunotherapy in B16F10 melanoma-bearing mice. FDG PET did not detect the response. Further studies are required to determine whether FBPA PET is useful in evaluating the treatment effect of ICIs in humans.

KEYWORDS

immune checkpoint inhibitor, melanoma, mouse, FBPA, FDG, PET

Introduction

Cancer immunotherapy with immune checkpoint inhibitors (ICIs) has increasingly been recognized as a novel effective treatment recently.

However, the treatment response is unsatisfactory for most patients as the response rate is limited to about 20% - 40% (1, 2). Methods to predict or determine the efficacy at an early stage of treatment are highly desired due to its high cost and possible autoimmune-like side effects named as immune-related adverse events (irAEs).

Positron emission tomography (PET), a functional and metabolic imaging technique, is a promising candidate of the assessment method in view of evaluating the disease activity of primary and metastatic malignant lesions. PET using [^{18}F] fluorodeoxyglucose (FDG PET) has widely been used for tumor imaging since glucose metabolism is enhanced in various types of malignancies. However, FDG PET may not be suitable to determine the effect of cancer immunotherapy at an early stage of treatment because of the false-positive findings caused by inflammation. Theoretically, cancer immunotherapy is accompanied by inflammation in the treatment area and glucose metabolism is enhanced by the inflammation in addition to tumor activity, which is visualized as increased FDG uptake. irAEs caused by ICIs include inflammation in various organs as well. Mekki, et al. reported that thoracic sarcoid-like reaction, enterocolitis, thyroiditis, hypophysitis, and pancreatitis were observed as irAEs in FDG PET (3). These inflammatory changes potentially provide false-positive FDG PET findings.

PET with L-4-borono-2- [^{18}F] fluoro-phenylalanine (FBPA), an amino acid-based radiotracer, has been used for pretreatment assessment before boron neutron capture therapy (BNCT) for cancer (4). BNCT is a type of radiotherapy based on the nuclear reaction of [^{10}B] (n, α) [^7Li]; a neutron beam from a nuclear

reactor or accelerator is irradiated around the ^{10}B containing tumor target and the emitted alpha particles have a higher cytotoxic effect and shorter range than beta rays. L-paraboronophenylalanine (BPA) labeled with ^{10}B is the major carrier compound used to deliver the boron selectively to the tumor cells (5). A recent study demonstrated that BPA was delivered to the cells through transporter-mediated mechanisms and that L-type amino acid transporter 1 (LAT1) was the major amino acid transporter related to these mechanisms (6). Another recent study reported that FBPA accumulated into tumor cells mainly *via* LAT1 and that FBPA uptake was significantly lower than FDG uptake in inflammatory lesions (7). FBPA is considered as a promising tumor-specific PET tracer.

In the present study, we aimed to demonstrate that FBPA PET was useful for evaluating the early response of anti-PD-1 cancer immunotherapy in B16F10 melanoma-bearing mice, comparing the results with those of FDG PET. An increase in FDG uptake was shown with anti-PD-1 treatment in a previous study using the same experimental model as this study (8).

Materials and methods

Animal model

C57BL/6JmsSlc mice (6 - 7 weeks old) were purchased from Japan SLC, Inc. (Hamamatsu, Japan). B16F10 melanoma cell line was obtained from ATCC (Manassas, VA, USA). Cells were cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin, and 2 mmol/L L-glutamine.

The mice were inoculated subcutaneously with 3×10^6 B16F10 cells in 100 μL of phosphate-buffered saline (PBS) 3 days before pretreatment PET imaging on Day 0.

Anti-PD-1 treatment

Anti-mouse PD-1 antibody (clone RMP1-14) was purchased from BioXcell (West Lebanon, NH, USA). The antibody (250 µg) in 200 µL PBS was administered twice intraperitoneally to the tumor-bearing mice on Day 0 and Day 5. Mice with anti-PD-1 antibody were defined as the treatment group. In addition to the treatment group, the control group was prepared, which consisted of mice receiving 200 µL PBS without anti-PD-1 antibody twice intraperitoneally on Day 0 and Day 5.

Tumors were measured using calipers, and the tumor volume expressed in mm³ was calculated according to the following formula: $0.5 \times (\text{long diameter}) \times (\text{short diameter})^2$. Relative tumor volume, the volume ratio of posttreatment to pretreatment tumor, was also calculated in this study.

Synthesis of [¹⁸F]FBPA

[¹⁸F]FBPA was synthesized according to previous report (9). In short, [¹⁸F]FBPA was produced by reacting a precursor solution (30 mg of 4-borono-L-phenylalanine in 4 mL trifluoroacetic acid) with [¹⁸F]acetylhyposulfonate. After the reaction, [¹⁸F]FBPA was separated by using High-Performance Liquid Chromatography (HPLC), and the HPLC solvent was removed before [¹⁸F]FBPA was recovered with saline. The radiochemical purity of [¹⁸F]FBPA was >98% and the molar activity was 232 GBq/mmol on average.

PET imaging

PET/computed tomography (PET/CT) imaging was performed twice for each mouse on Day 0 before the anti-PD-1 antibody administration and on Day 7 using a micro-PET/CT scanner (Inveon, Siemens, Munich, Germany).

FBPA or FDG PET/CT studies were conducted separately in 12 tumor-bearing mice (n = 6 each for the treatment and control groups).

FBPA or FDG was injected *via* the tail vein of each mouse (FBPA: 2.18 ± 0.34 MBq, FDG: 2.17 ± 0.25 MBq). Static PET images were acquired for 10 min starting at 60 min after radiotracer injection under isoflurane anesthesia.

PET images were reconstructed using three-dimensional ordered-subset expectation-maximization algorithm (16 subsets, 2 iterations) with attenuation and scatter correction.

Image analysis

The radioactivity of each tumor was expressed quantitatively as the standardized uptake value (SUV), which was corrected for

the injected dose (MBq) and body weight (g). The maximum value of SUV (SUV_{max}) was calculated from a single voxel exhibiting the maximum SUV in each tumor. The mean of SUV (SUV_{mean}) was obtained in the liver as a background value.

Spherical volumes of interest (VOIs) were placed on the tumor and liver in the PET images using commercially available software (PETSTAT, AdIn Research, Tokyo, Japan) while referring to the CT images.

The tumor to liver ratio (T/L ratio) was also used as a parameter exhibiting tumor activity, which was defined by the ratio of SUV_{max} in the tumor to SUV_{mean} in the liver. In addition to pre- or posttreatment parameters in the tumor, changes between pre- and posttreatment tumor activity were observed using the formula as follows: $[(\text{posttreatment parameter values}/\text{pretreatment values} - 1) \times 100]$ (%). They were expressed as %SUV_{max} and %T/L ratio.

Tumor uptake volume (TUV) and metabolic tumor volume (MTV) were also calculated for FBPA and FDG, respectively. TUV or MTV was defined as the volume within a tumor margin, which was delineated with 40% of SUV_{max}. These quantitative parameters were also obtained from the same software mentioned above.

Immunohistochemistry

After the mice were sacrificed by euthanasia, tumor xenografts were resected and subjected to immunohistochemical staining. Immunohistochemical staining was performed to determine if LAT-1 and GLUT-1 transporters, respectively, for FBPA and FDG uptake changed between the treatment and control groups. CD8 and PD-1 proteins were also evaluated in the two groups to observe their changes after treatment.

The antibodies used in this study were anti-LAT1 antibody (orb96302, Biorbyt, Cambridge, UK) for FBPA, anti-GLUT-1 antibody (ab115730, abcam, Cambridge, UK) for FDG, anti-PD-1 antibody (#84651, CST, Danvers, MA, USA) for PD-1 expression, and anti-CD8 antibody (14-0808-80, Invitrogen, Waltham, MA, USA) for CD8 expression in accordance with the manufacturer's instructions. PD-1 and CD8 expressions were evaluated to recognize the association between these markers and FBPA or FDG uptake.

Serial 4-µm tumor paraffin-embedded sections were used for the immunohistochemical staining as well as hematoxylin-eosin (HE) staining.

Statistical analysis

Unpaired one-tailed t-tests were used to compare the values between the two groups. Statistical significance was set at $p < 0.05$.

Results

Tumor volume

Tumors on Day 1 represented no statistical difference in volume between the treatment and control groups. Tumor volumes in both groups also showed no statistical differences on Days 3 and 5. However, tumors in the treatment group were statistically smaller than those in the control group on Day 7 ($p < 0.01$, Figure 1A).

Relative tumor volume exhibited the same trend from Day 1 to Day 7, although no statistical difference was observed between the treatment and control groups on Day 7 (Figure 1B).

PET visual analysis

Both FBPA and FDG PET exhibited faint uptake in the tumor before treatment on Day 0 (Figures 2A: FBPA, 2B: FDG).

FBPA and FDG PET showed intense tumor uptake either in the treatment or control group on Day 7 (Figures 2C: FBPA, 2D: FDG). FBPA or FDG uptake was higher on Day 7 than Day 0 in each tumor.

Visual analysis solely did not allow differentiation of the tumors between the treatment and control groups.

PET quantitative analysis

FBPA and FDG PET before treatment

Both FBPA and FDG PET demonstrated similar SUVmax and T/L ratios of the tumors in the treatment and control groups before treatment on Day 0 without statistical differences. Mean

SUVmax and T/L ratio were 0.94 and 0.85 in the treatment group and 0.99 and 0.85 in the control group of FBPA PET, respectively. Mean SUVmax and T/L ratio were 1.0 and 1.7 in the treatment group and 1.1 and 1.4 in the control group of FDG PET, respectively.

FBPA PET after treatment

Posttreatment PET on Day 7 demonstrated that SUVmax, T/L ratio, and TUV in the treatment group were statistically smaller than those in the control group (Table 1, Figures 3A: SUVmax, 3B: T/L ratio 3C: TUV). Mean SUVmax, T/L ratio and TUV were 2.6, 2.3, and 0.18 in the treatment group and 3.2, 3.1, and 0.53 in the control group, respectively. %SUVmax and %T/L ratio in the treatment group tended to be smaller than those in the control group without statistical significances (Table 1). Mean %SUVmax and %T/L ratio were 183 and 178 in the treatment group and 225 and 267 in the control group, respectively.

FDG PET after treatment

Tumors in both groups exhibited no statistical differences in all parameters again after treatment on Day 7 (Table 1 and Figure 3A: SUVmax, 3B: T/L ratio, 3C: MTV). However, T/L ratio and %SUVmax in the treatment group represented larger values than those in the control group in this setting (Table 1). Mean SUVmax, T/L ratio and MTV were 5.1, 7.3, and 0.61 in the treatment group and 5.1, 5.4, and 1.0 in the control group, respectively. Mean %SUVmax and %T/L ratio were 403 and 341 in the treatment group and 331 and 305 in the control group, respectively.

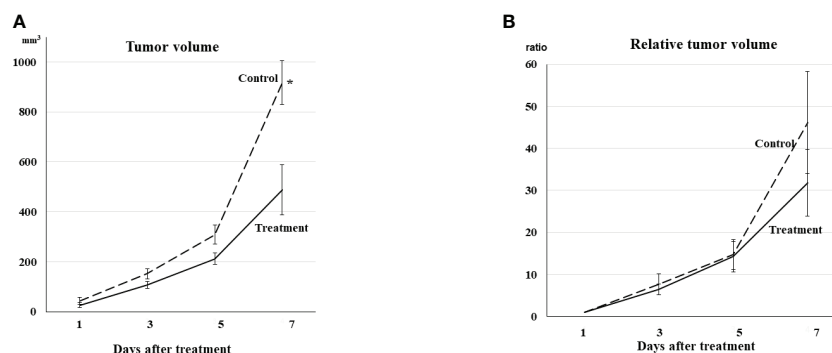


FIGURE 1

(A) Tumor volume Tumor volume increased gradually even in the treatment group. Tumor volumes in both groups showed no differences on Days 3 and 5. The tumors in the treatment group were statistically smaller than those in the control group on Day 7 ($*p < 0.01$). (B) Relative tumor volume Relative tumor volume to the volume on Day 1 exhibited the same trend from Day 1 to Day 7 as actual tumor volume, although no statistical difference was observed between the treatment and control groups on Day 7.

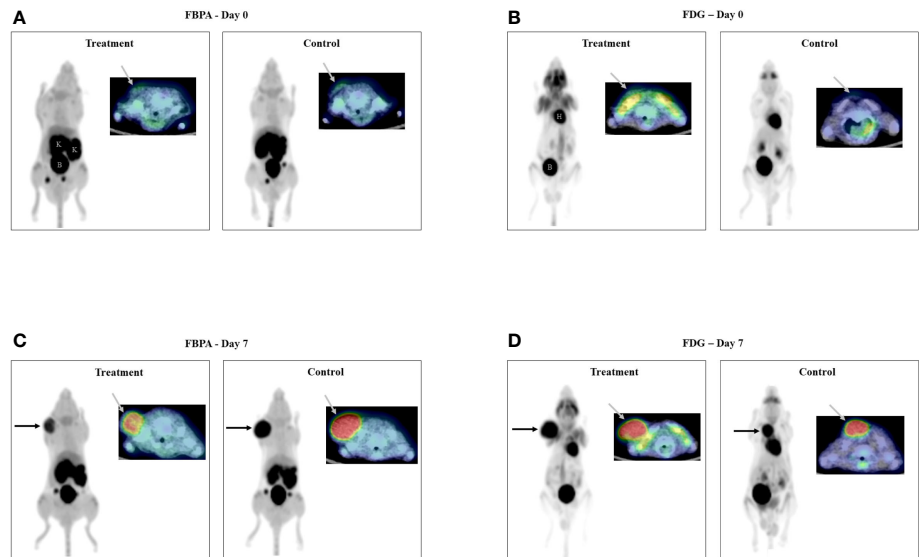


FIGURE 2
PET/CT images on Day 0 (2A: FBPA, 2B: FDG) Faint uptake (gray arrow) was observed in all tumors only on transaxial FBPA and FDG PET/CT images before treatment. PET/CT images on Day 7 (2C: FBPA, 2D: FDG) Intense uptake (black and gray arrows) was observed in all tumors on FBPA and FDG PET/CT images. Visual analysis did not discriminate the tumors in the treatment and control groups either on FBPA or FDG PET/CT images. Radioactivities were observed in the kidneys (K) and bladder (B) and in the heart (H) and bladder (B) in addition to the tumor, respectively, as shown in FBPA and FDG PET images on Day 0. (Left: maximum intensity projection image of PET, right upper: fused PET/CT image, Treatment: PET image in the treatment group, Control: PET image in the control group).

TABLE 1 Comparison of quantitative parameters between the treatment and control groups in PET using FBPA or FDG.

	Treatment	FBPA Control		Treatment	FDG Control	
SUVmax	2.6 ± 0.22	3.2 ± 0.2	p <0.05	5.1 ± 0.82	5.1 ± 1.1	n.s.
T/L ratio	2.3 ± 0.26	3.1 ± 0.28	p <0.05	7.3 ± 0.62	5.4 ± 1.4	n.s.
%SUVmax	183 ± 32	225 ± 24	n.s.	403 ± 43	331 ± 65	n.s.
%T/L ratio	178 ± 32	267 ± 38	n.s.	341 ± 33	305 ± 73	n.s.
TUV or MTV	0.18 ± 0.061	0.53 ± 0.035	p <0.005	0.61 ± 0.24	1.0 ± 0.15	n.s.

TUV, Tumor uptake volume. n.s., not significant.
n.s., not significant.

Immunohistochemical analysis

Intense LAT1 and GLUT-1 expression, respectively, for FBPA and FDG uptake was observed in many areas within the tumor, whereas PD-1 and CD8 expression was in the limited areas within the tumor in both the treatment and control groups. Visual analysis did not discriminate the expression of LAT1, GLUT-1, PD-1, or CD8 between the treatment and control groups (Figure 4). These results were in line with the PET findings that intense FBPA and FDG tumor uptake was observed either in the treatment or control group on Day 7. Treatment-induced inflammation was not prominent in the model used in this study.

Discussion

FBPA and FDG were used as radiotracers of PET imaging in this experimental study and discordant findings were observed between them at an early stage of the anti-PD 1 immunotherapy. Although conducted in the tumor-bearing mice, this study firstly compared an amino acid-based radiotracer and FDG in such an immunotherapy setting. The promising findings of FBPA PET in this study warrant further studies of this imaging technique in evaluating the response of immunotherapy with ICIs in humans. Tumors in the treatment group were smaller than those in the control group on Day 7 in this study. However, all tumors in the treatment and control groups were on the way of growing,

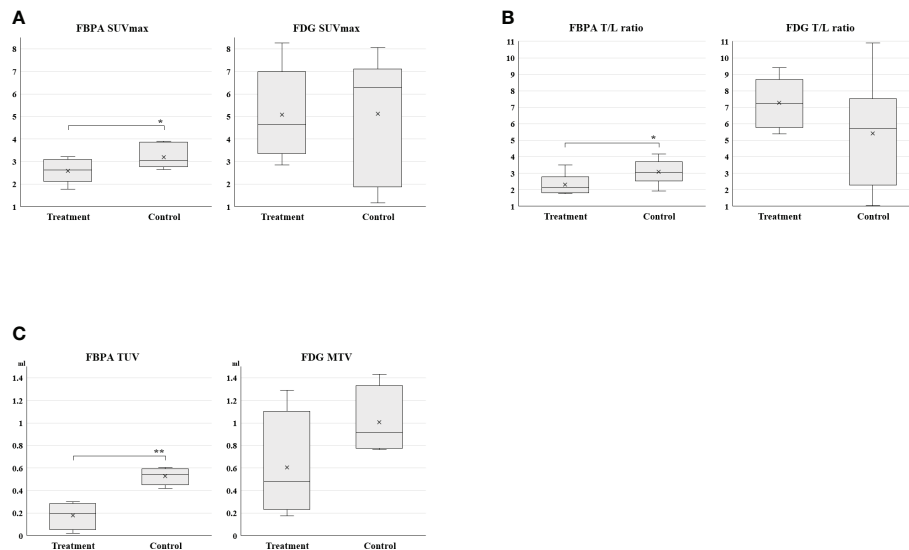


FIGURE 3

(3A: SUVmax, 3B: T/L ratio, 3C: TUV and MTV) Box charts of quantitative PET parameters in the treatment and control groups. Statistically smaller values of SUVmax, T/L ratio, and TUV were observed in the treatment group than the control group on FBPA PET (* $p < 0.05$, ** $p < 0.005$). A substantial overlap was observed in these parameters on FDG PET.

and size reduction of the tumors was not observed even in the treatment group. A similar trend of tumor growth was reported in the previous studies dealing with anti-PD-1 treatment in the same tumor-bearing mouse model (8, 10). The slightly smaller tumor volume in the treatment group compared to the control group was considered a reflection of the early treatment effect with the anti-PD-1 antibody in mice with B16F10 melanoma.

FBPA allowed detection of the anti-PD-1 treatment effect on Day 7. SUVmax, T/L ratio, and TUV in the treatment group represented smaller values than those in the control group with statistical significances. %SUVmax and %T/L ratio exhibited a similar trend although statistical significances were not observed. FBPA has been reported to be a tumor-specific PET tracer and shows low uptake in inflammatory lesions (7). Additionally, this study demonstrated that FBPA was useful in evaluating the early response by the anti-PD-1 therapy, which is known to cause immune and inflammatory reactions in the tumor.

Only a limited number of studies have been reported so far regarding the use of amino acid-based PET radiotracers in evaluating the treatment response to cancer immunotherapy with ICIs. Galldiks, et al. reported the additional value of ^{18}F FET <(O-(2-[^{18}F]fluoroethyl)-L-tyrosine)> to contrast-enhanced MRI for treatment monitoring of immunotherapy with ICIs or targeted therapy (TT) alone or in combination with radiotherapy in patients with metastatic brain tumors (11). FET PET seemed to be of great value for the differentiation of treatment-related changes from metastatic brain tumors.

Tomita, et al. reported that anti-PD-1 treatment increased mean FDG uptake values in the tumor in the same tumor-

bearing mouse model as this study (8). Maximum FDG uptake values in the tumor represented no statistical difference in their study. In this study, quantitative parameters such as SUVmax, T/L ratio, and MTV in the tumors exhibited no statistical differences between the treatment and control groups. However, T/L ratio and %SUVmax in the treatment group tended to represent larger values than those in the control group. The slight increase in these quantitative parameters appeared to correspond to the increased tumor FDG uptake observed in the study of Tomita, et al. as a reflection of the immune response caused by the anti-PD-1 treatment.

We used SUVmax and T/L ratio to express radiotracer uptake in the tumor in PET quantitative analysis. Although SUVmax or SUVmean has frequently been used as a quantitative parameter in research using PET, it requires the injected dose of radiotracer for calculation. The estimation of actual injected doses is difficult in small animals as the residual radioactivity in tail veins or syringes after injection is relatively large compared to the radioactivity of doses prepared. In this regard, T/L ratio might be more reliable than SUVmax since it is a completely image-derived parameter.

Volumetric FDG PET parameters, MTV and TLG, are known to be better than SUVmax in evaluating or predicting chemotherapeutic responses in clinical situations. Recent studies also demonstrated that volumetric parameters were useful in evaluating the early response by the immunotherapy with ICIs (12, 13). The volumetric parameter, TUV, was also used for FBPA PET in this study. TUV in addition to SUVmax and T/L ratio in the treatment group represented smaller values than those in the control group. However, MTV in

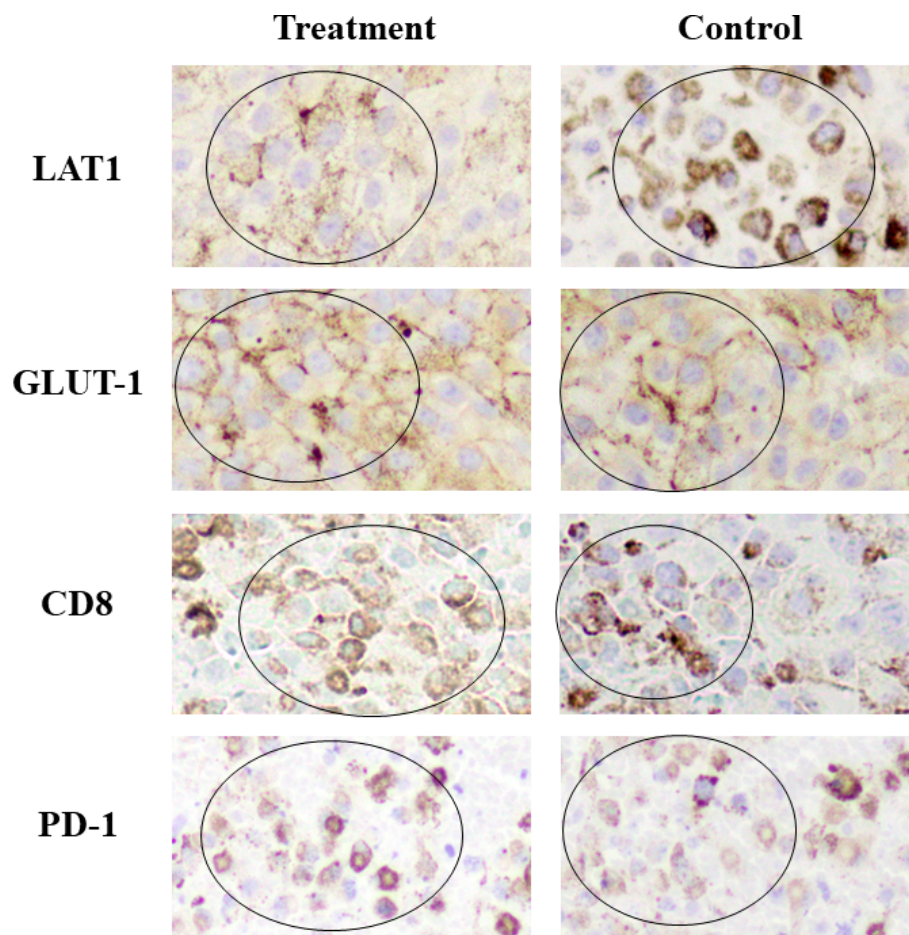


FIGURE 4

Immunohistochemical staining Intense LAT1 and GLUT-1 expression was observed in many areas within the tumor, whereas CD8 and PD-1 expression was in the limited areas within the tumor in both the treatment and control groups. (Positive staining: mainly observed in circled areas) No differences in the expression of LAT1, GLUT-1, CD8, or PD-1 were observed visually between the treatment and control groups. GLUT-1 and CD8 expression in the treatment group appeared slightly higher than that in the control group.

FDG PET exhibited no statistical differences between the treatment and control groups. This study successfully demonstrated that volumetric PET parameters were useful in evaluating the early response of anti-PD-1 treatment in the tumor-bearing mice.

IHC analysis demonstrated that intense LAT1 or GLUT-1 expression was observed in many areas within the tumor, whereas CD8 expression was in the limited areas. No obvious differences in these expressions were observed between the treatment and control groups. The clear mechanism of the slight increase in tumor FDG uptake after treatment was not resolved in this study. An experimental study using the same tumor-bearing mouse model as our study demonstrated that anti-PD-1 therapy increased glucose metabolism by cancer cells themselves at an early stage of treatment (8). The subtle inflammatory change caused by anti-PD-1 therapy was not considered to largely affect glucose metabolism in that study.

This study has some limitations. As we would like to demonstrate the advantages of FBPA over FDG in detecting early response of anti-PD-1 immunotherapy, we compared these two radiotracers in the same experimental model as used in the paper dealing with FDG uptake (8). Thus, we used only one treatment protocol with early drug administration, low-dose and short-term regimen. We only used a B16F10 melanoma tumor-bearing mouse model according to the previous studies dealing with anti-PD-1 treatment (8, 10). Thus, the results in this study may not apply to other kinds of tumors in mice or in humans. The observation of tumor growth was limited up to a few days from PET imaging due to an ethical consideration to small animals. An obvious reduction in tumor size was not confirmed in both the treatment and control groups. However, smaller tumor volume in the treatment group compared to the control group was observed and was considered a reflection of the early treatment effect with the anti-PD-1 treatment as stated above. Although discrepant

quantitative findings were observed after treatment between FBPA and FDG PET, these findings were obtained from tumors in different mice. Direct comparison of the FBPA and FDG findings in the same tumor would have been ideal, but it was impossible as both FBPA and FDG are ^{18}F labeled radiopharmaceuticals.

Conclusion

This study demonstrated that FBPA PET allowed detection of the early response of anti PD-1 immunotherapy in B16F10 melanoma-bearing mice. FDG PET did not detect the response. Further studies are required to determine whether FBPA PET is useful in evaluating treatment effect of ICIs in humans.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The animal study was reviewed and approved by Animal Care and Use Committee of the Osaka University Graduate School of Medicine.

Author contributions

The study was designed by MT and FS. All authors participated in collecting data. MT and HF prepared the

manuscript and contributed to data analysis and interpretation. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Japan Society for the Promotion of Science KAKENHI Grant Number 19K08198.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.1026608/full#supplementary-material>

References

- Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* (2017) 168:707–23. doi: 10.1016/j.cell.2017.01.017
- Pons-Tostivint E, Latouche A, Vaflard P, Ricci F, Loirat D, Hescot S, et al. Comparative analysis of durable responses on immune checkpoint inhibitors versus other systemic therapies: A pooled analysis of phase III trials. *JCO Precis Oncol* (2019), 3:1–10. doi: 10.1200/po.18.00114
- Mekki A, Derclé L, Lichtenstein P, Marabelle A, Michot JM, Lambotte O, et al. Detection of immune-related adverse events by medical imaging in patients treated with anti-programmed cell death 1. *Eur J Cancer* (2018) 96:91–104. doi: 10.1016/j.ejca.2018.03.006
- Imahori Y, Ueda S, Ohmori Y, Kusuki T, Ono K, Fujii R, et al. Fluorine-18-labeled fluoroboronophenylalanine PET in patients with glioma. *J Nucl Med* (1998) 39:325–33.
- Pisarev MA, Dagrosa MA, Juvenal GJ. Boron neutron capture therapy in cancer: past, present and future. *Arq Bras Endocrinol Metabol* (2007) 51:852–6. doi: 10.1590/s0004-27302007000500024
- Wongthai P, Hagiwara K, Miyoshi Y, Wiriyaerkmul P, Wei L, Ohgaki R, et al. Boronophenylalanine, a boron delivery agent for boron neutron capture therapy, is transported by ATB0+, LAT1 and LAT2. *Cancer Sci* (2015) 106:279–86. doi: 10.1111/cas.12602
- Watabe T, Ikeda H, Nagamori S, Wiriyaerkmul P, Tanaka Y, Naka S, et al. 18F-FBPA as a tumor-specific probe of L-type amino acid transporter 1 (LAT1): a comparison study with 18F-FDG and 11C-methionine PET. *Eur J Nucl Med Mol Imaging* (2017) 44:321–31. doi: 10.1007/s00259-016-3487-1
- Tomita M, Yasui H, Higashikawa K, Nakajima K, Takakura H, Shiga T, et al. Anti PD-1 treatment increases [18F]FDG uptake by cancer cells in a mouse B16F10 melanoma model. *EJNMMI Res* (2018) 8:1–8. doi: 10.1186/s13550-018-0433-1
- Naka S, Watanabe T, Kanai Y, Watabe T, Tatsumi M, Kato H, et al. Improved stability and practicality for synthesis of 4-Borono-2-[18F]fluoro-L-phenylalanine by combination of [18O]O₂ single-use and [18F]CH₃COOF labeling agents. *Nucl Med Mol Imaging* (2010) (2022) 56:86–95. doi: 10.1007/s13139-021-00719-1
- Chen S, Lee LF, Fisher TS, Jessen B, Elliott M, Evering W, et al. Combination of 4-1BB agonist and PD-1 antagonist promotes antitumor effector/memory CD8 T cells in a poorly immunogenic tumor model. *Cancer Immunol Res* (2015) 3:149–60. doi: 10.1158/2326-6066.CIR-14-0118
- Gallidiks N, Abdulla DSY, Scheffler M, Wolpert F, Werner JM, Hullner M, et al. Treatment monitoring of immunotherapy and targeted therapy using 18F-FET PET in patients with melanoma and lung cancer brain metastases: Initial experiences. *J Nucl Med* (2021) 62:464–70. doi: 10.2967/jnumed.120.248278

12. Umeda Y, Morikawa M, Anzai M, Ameshima S, Kadowaki M, Waseda Y, et al. Predictive value of integrated 18 F-FDG PET/MRI in the early response to nivolumab in patients with previously treated non-small cell lung cancer. *J Immunother Cancer* (2020) 8:1–11. doi: 10.1136/jitc-2019-000349

13. Voltin CA, Mettler J, van Heek L, Goergen H, Muller H, Baues C, et al. Early response to first-line anti-PD-1 treatment in hodgkin lymphoma: A PET-based analysis from the prospective, randomized phase II NIVAHL trial. *Clin Cancer Res* (2021) 27:402–7. doi: 10.1158/1078-0432.CCR-20-3303



OPEN ACCESS

EDITED BY
Ira Blevis,
Philips Healthcare, United States

REVIEWED BY
Michael Götz,
University of Ulm, Germany
Ali Selamat,
University of Technology Malaysia,
Malaysia
Shujaat Khan,
Siemens Healthineers, United States

*CORRESPONDENCE
Lidia Garrucho
✉ lgarrucho@ub.edu

SPECIALTY SECTION
This article was submitted to
Breast Cancer,
a section of the journal
Frontiers in Oncology

RECEIVED 14 September 2022

ACCEPTED 19 December 2022

PUBLISHED 23 January 2023

CITATION
Garrucho L, Kushibar K, Osuala R,
Diaz O, Catanese A, del Riego J,
Bobowicz M, Strand F, Igual L and
Lekadir K (2023) High-resolution
synthesis of high-density breast
mammograms: Application to
improved fairness in deep learning
based mass detection.
Front. Oncol. 12:1044496.
doi: 10.3389/fonc.2022.1044496

COPYRIGHT
© 2023 Garrucho, Kushibar, Osuala,
Diaz, Catanese, del Riego, Bobowicz,
Strand, Igual and Lekadir. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

High-resolution synthesis of high-density breast mammograms: Application to improved fairness in deep learning based mass detection

Lidia Garrucho^{1*}, Kaiser Kushibar¹, Richard Osuala¹,
Oliver Diaz¹, Alessandro Catanese², Javier del Riego³,
Maciej Bobowicz⁴, Fredrik Strand⁵, Laura Igual¹
and Karim Lekadir¹

¹Barcelona Artificial Intelligence in Medicine Lab, Facultat de Matemàtiques i Informàtica, Universitat de Barcelona, Barcelona, Spain, ²Unitat de Diagnòstic per la Imatge de la Mama (UDIM), Hospital Germans Trias i Pujol, Badalona, Spain, ³Àrea de Radiologia Mamària y Ginecològica (UDIAT CD), Parc Taulí Hospital Universitari, Sabadell, Spain, ⁴2nd Department of Radiology, Medical University of Gdansk, Gdansk, Poland, ⁵Breast Radiology, Karolinska University Hospital and Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden

Computer-aided detection systems based on deep learning have shown good performance in breast cancer detection. However, high-density breasts show poorer detection performance since dense tissues can mask or even simulate masses. Therefore, the sensitivity of mammography for breast cancer detection can be reduced by more than 20% in dense breasts. Additionally, extremely dense cases reported an increased risk of cancer compared to low-density breasts. This study aims to improve the mass detection performance in high-density breasts using synthetic high-density full-field digital mammograms (FFDM) as data augmentation during breast mass detection model training. To this end, a total of five cycle-consistent GAN (CycleGAN) models using three FFDM datasets were trained for low-to-high-density image translation in high-resolution mammograms. The training images were split by breast density BI-RADS categories, being BI-RADS A almost entirely fatty and BI-RADS D extremely dense breasts. Our results showed that the proposed data augmentation technique improved the sensitivity and precision of mass detection in models trained with small datasets and improved the domain generalization of the models trained with large databases. In addition, the clinical realism of the synthetic images was evaluated in a reader study involving two expert radiologists and one surgical oncologist.

KEYWORDS

data synthesis, full-field digital mammograms, generative adversarial networks (GANs), data augmentation (DA), mass detection, reader study, breast cancer

1 Introduction

Breast density is divided into four categories in the American College of Radiology Breast Imaging and Data System (ACR BI-RADS) 5th edition (1). The categories range from A to D and correspond to fatty, scattered, heterogeneous, and extremely dense breasts. The qualitative classification of breast density in mammography is an accepted method in breast radiology with good inter-observer and intra-observer agreement (2), despite the fact that commercial software can produce a more accurate quantitative measure by calculating the ratio of fibroglandular tissue to the total breast area.

In mammography databases, the distribution of breast densities among women aged 40 years or older is approximately 43% for dense breasts: 36% for BI-RADS C and 7% for BI-RADS D (3). Han et al. (4) found that women with a family history of breast cancer were more likely to have dense breasts than women with no cancer in the family history. In addition, high breast density is associated with an increased risk of interval cancers (5), being those 13–31 times more likely in BI-RADS D breasts than in BI-RADS A (6–8). Consequently, it is recommended to decrease the interval between screening mammograms and consider supplemental screening for women with dense breasts (8, 9).

Dense breast tissue is one of the strongest and most common independent risk factors for breast cancer (5, 6, 10). On mammograms, masses and other suspicious findings can be obscured in normal dense tissue and become imperceptible on mammograms. Therefore, the sensitivity of mammography decreases with increasing breast density and has a range value of 81–93% for fatty breasts, 84–90% for breasts with scattered fibroglandular density, 69–81% for heterogeneously dense breasts, and 57–71% for extremely dense breasts in women 40–74 years of age (5). Although mammography is the gold standard non-invasive method for breast cancer detection in population-

based screening, women with dense breasts have shown both a reduced cancer detection and higher mortality rates (11–13).

The goal of this study is two-fold. First, mitigate the differences in computer-aided detection (CADe) systems sensitivity by breast density (14). Second, improve the performance of state-of-the-art deep learning-based breast mass detection models by the means of synthetic data augmentation.

Figure 1 shows the differences in sensitivity of our deep learning-based mass detection model by density composition. The breast composition distribution of the dataset used to train the model has a big unbalance in categories A (9%) and D (5%) (Figure 1A). Nonetheless, the sensitivity between fatty (95% for BI-RADS A) and extremely dense breasts (82% for BI-RADS D) is very different (Figure 1B). The decrease in performance is partly caused by the high rate of false positives in extremely dense breasts (Figure 1C).

Data augmentation is commonly used to increase the variability of the training samples, improve the model generalization and avoid overfitting. Among all deep learning-based augmentation techniques, Generative Adversarial Networks (GANs) (15) are frequently used to generate new synthetic samples in an unsupervised manner. GANs have been previously used to synthesize full-field digital mammograms (FFDMs) or lesion patches, normally at low resolutions (16–18). Becker et al. (19) trained a cycle-consistent GAN (CycleGAN) on downsampled mammograms (256×256 and 515×408 pixels) to artificially inject or remove suspicious features. In their reader study, three radiologists could discriminate between original and synthetic images with an area under the curve (AUC) of 0.94, mainly due to the presence of artifacts. Zakka et al. (20) trained a style-based GAN to generate 512×512 mammograms enabling user-controlled global and local attribute-editing. Then, a double-blind study involving four expert radiologists assessed the quality of the resulting images achieving an average AUC of 0.54.

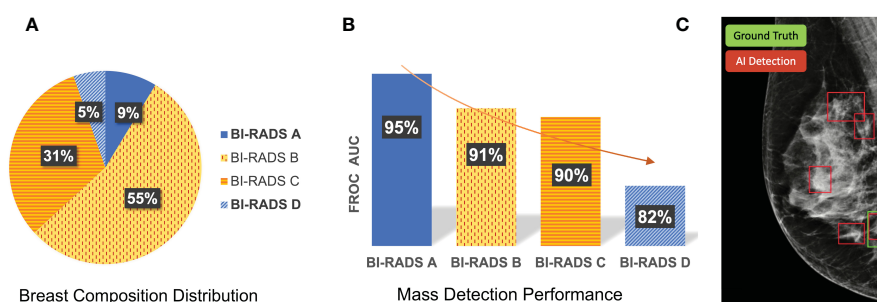


FIGURE 1

Differences in sensitivity of a deep learning-based mass detection model by breast density composition in full-field digital mammography. (A) Breast composition distribution of the training set. (B) Mass detection sensitivity by breast density. (C) Output of an automated detection model in a high-density (BI-RADS D) image. The red bounding boxes are the AI model outputs and the green box corresponds to the true mass, summing a total of four false positives.

Other studies used the synthetic mammograms as data augmentation to improve the performance in different downstream tasks. Synthetic data augmentation using GANs was evaluated in breast cancer classification by Shrinivas et al. (21). The proposed model, a Deep Convolutional GAN (DCGAN), synthesized FFDM with 256×256 image resolution. The synthetic images were validated using a Visual Turing test with the help of medical experts and were easily spotted by the radiologist because they lacked the sharpness and fine-grained details of original mammograms. In a similar study, Jendele et al. (22) balanced the ratio of benign and malignant lesions in the training set using a CycleGAN trained to translate healthy mammograms to mammograms containing malignant findings. The synthetic mammograms were 256×256 pixels, as higher image resolutions introduced many artifacts. The benefits of using the synthesized mammograms for data augmentation were inconclusive considering that the performance of their detection model did not improve. Muramatsu et al. (23) trained a CycleGAN using masses from unrelated domains – lung CT and mammography – to synthesize 256×256 pixels masses and improve the mass classification in mammography. However, no statistical difference was found between the model trained with synthetic masses and the classifier trained with original mass patches.

In this study, the original resolution of FFDMs, around 5 Megapixels (MP), with image sizes of 3328×4084 or 2560×3328 pixels depending on the manufacturer. Two main challenges have prevented the use of GANs for high-resolution FFDM synthesis. The first one is the high demand for graphics processing unit (GPU) memory, which typically scales with the input and output resolutions. As an example, CycleGAN needs more than 24GB of GPU memory when the input image is larger than 1MP. The second challenge is data scarcity. The training set has to be representative enough to generate realistic samples and overcome the training instabilities and image artifacts of GANs. High intrinsic heterogeneity exists across mammograms due to the huge variability of breast sizes, shapes, and compositions. Moreover, FFDMs contain very fine structural details at high resolution such as the different parenchymal patterns, nipples, and pectoral muscles, the presence of lymph nodules, microcalcifications, or calcifications, among many other associated features.

Korkinof et al. (24) was the first study that managed to synthesize high-resolution FFDM images using a Progressively Growing GAN (PGGAN) (25). Their PGGAN was trained using more than 400,000 FFDMs and was demonstrated to generate mammograms up to a 1280×1024 pixel resolution. All mammograms available were used for training, independently of the clinical findings, and only images with post-operative artifacts and large foreign bodies such as implants were excluded from the training set. In a separate retrospective study, Kornikof et al. (26) evaluated the perceived realism of the synthetic FFDM images in a reader study involving 55 radiologists and 62 non-

radiologists. Overall, in the setup of this study, the synthetic images were shown to be indistinguishable from original mammograms. However, it was unclear and was not further investigated whether the synthetic FFDMs have relevant applications for clinical purposes.

This work presents for the first time the use of GAN models to generate high-resolution FFDMs with increased breast density using images from different manufacturers and datasets. Moreover, we evaluated the potential of using the synthesized images as data augmentation to improve the mass detection performance. Only a single prior study had performed similar data augmentation by breast density categories for improved mass detection (27). However, the authors employed mathematical breast phantoms generated using the pipeline in the VICTRE study (28) instead of GANs. The breast phantoms were generated across the four BI-RADS breast density categories for each view, cranio-caudal (CC) and mediolateral oblique (MLO), and were modeled after a single vendor, the Siemens Mammomat Inspiration. The limitation of their study is the lack of diversity within each density type, including the size and shape of the breast. Moreover, the statistical analysis did not show a significant difference between the Free-response Receiver Operating Characteristic (FROC) curves for mass detection.

To summarize, the contributions of our study are as follows:

- Synthesize high-resolution high-density FFDMs using GAN-based models from three different datasets and mammography systems (manufacturers).
- Tackle the class imbalance by breast density composition by augmenting the training set using high-density synthetic mammograms.
- Improve the performance of mass detection in extremely dense breasts, categorized as BI-RADS D.
- Investigate the potential of high-density data augmentation for domain adaptation.
- Evaluate the anatomical realism of the synthetic mammograms in a reader study involving two expert radiologists and one surgical oncologist.

2 Materials and methods

2.1 Datasets and breast density

Four different datasets were used in this study. General details of these datasets are presented in Table 1. The OPTIMAM Mammography Image Database (OMI-DB) (29) comprises FFDM from the Breast Screening Programme of the United Kingdom (UK). In this study, we used the subset of mammograms captured with the Hologic Selenia Dimensions scanners (Hologic, Inc., Massachusetts, United States). The proportion of fibroglandular (dense) tissue in the breast was

TABLE 1 Digital mammography datasets used in this study.

	BCDR		CSAW		OPTIMAM Hologic			INbreast	
	ACR	Normal MMG	LIBRA (%)	Normal MMG	Volpara VBD(%)	Normal MMG	MMG with masses	ACR	MMG with masses
BI-RADS A	1	40	≤ 2.8	435	≤ 3.5	972	344	1	42
BI-RADS B	2	40	(2.8, 25)	52064	(3.5, 7.5)	3670	1740	2	36
BI-RADS C	3	62	(25, 75)	38545	(7.5, 15.5)	1987	808	3	21
BI-RADS D	4	58	≥ 75	394	≥ 15.5	708	161	4	8

For each dataset, the total number of mammograms (MMG) with the available breast density information are mapped to the corresponding BI-RADS categories. In CSAW, the breast percentage was obtained running LIBRA software. In OPTIMAM Hologic, the Volumetric Breast Density (VBD) percentage was obtained using Volpara software. In BCDR and INbreast, only the American College of Radiology (ACR) categories are available in the dataset information.

obtained from the commercial Volpara software (version 1.5.4, Volpara Health, Wellington, New Zealand). The Volumetric Breast Density (VBD) percentage of each mammogram was mapped to the corresponding BI-RADS breast density category. Only normal mammograms – without pathologies – were selected to train the synthetic data generation models. The mammograms containing masses were used to train and test the mass detection AI model.

The non-hidden case-control dataset of the CSAW dataset (30) was also used in this study. The dataset comprises screening FFDMs from Karolinska University Hospital (Solna, Sweden) acquired with Hologic Inc. devices. A total of 91,484 normal FFDM contained breast density information estimated using the LIBRA automated tool (University of Pennsylvania, United States) (31). The mapping between LIBRA breast density percentage and BI-RADS A and D categories was done by selecting the tails of the percent density distribution of the healthy exams.

The Breast Cancer Digital Repository (BCDR) dataset (32) is a public dataset from 2012 comprising images supplied by the Centro Hospitalar São João, at University of Porto (Portugal) and obtained using a MammoNovation Siemens FFDM scanner. For our purposes, we selected a total of 200 FFDMs without pathologies. The breast density categories were provided in the annotations of the dataset following the American College of Radiology (ACR) statement on reporting breast density (1).

The INbreast dataset (33) was acquired from a single Portuguese center using a FFDM system, the Siemens MammoNovation. INbreast was used in this study as an external validation dataset for the mass detection model trained with OPTIMAM Hologic images. In addition, INbreast was used to train a mass detection model in a low data regime scenario. The dataset contains 107 FFDMs with 116 annotated masses from different breast densities. The percentage of images in each BI-RADS category is 36%, 35%, 22% and 7% for BI-RADS A, B, C and D.

All images have a matrix of 3328×4084 or 2560×3328 pixels, depending on the compression plate used for image acquisition. In a previous work, we confirmed that a resolution of 1332×800

pixels was enough to detect small masses and reach state-of-the-art performance in different AI detection methods (34). All the FFDM images were cropped to the breast region and resized to 1332×800 pixels keeping the aspect ratio. Our target resolution for data synthesis was the same as the one used by our deep learning mass detection model.

2.2 Synthesis of high-density full-field digital mammograms

Our goal was to synthesize high-density FFDM from original low density mammograms and then use the synthetic data to improve the performance of our mass detection models. To this end, the training images were split by breast density BI-RADS categories, being *BI-RADS A* the source domain and *BI-RADS D* the target domain. Before training, all mammogram images were resized to the target resolution (1332×800 pixels), the input size of the mass detection model.

The CycleGAN (35) was the method selected to perform the low-to-high-density mammogram translation. The choice of CycleGAN was motivated by its widespread use and successes reported in the cancer imaging domain (17). A key methodological feature of CycleGAN is that its training data can be unpaired without the need for corresponding image pairs in source and target domains. Unpaired training data ensured the applicability of CycleGAN to our datasets, in which image pairs are not available, as the same breast of the same patient cannot be from both the high and the low breast density domains.

As shown in Figure 2, the CycleGAN contains two mapping functions: 1) an image x in the source domain is mapped to a synthetic image $G(x) = \hat{y}$ in the target domain via a generator G ; and 2) an image y is mapped from target to source $F(y) = \hat{x}$ via a generator F . This enables translating images from source to target and back to the source domain $F(G(x)) = \hat{x}$. Both generators F and G are paired with corresponding discriminators, which try to classify whether a generated image is real or synthetic in a two-player *minmax* game with their respective generator (15). Based on the predictions of the

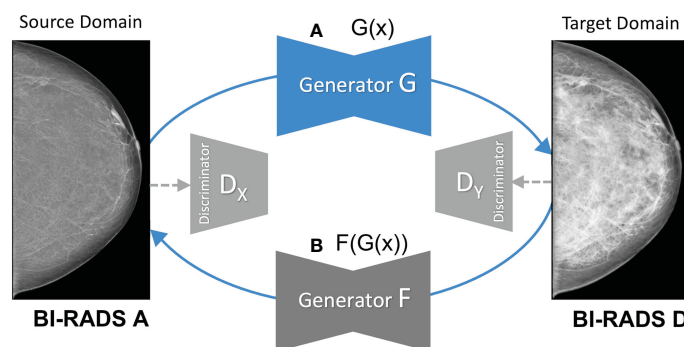


FIGURE 2

Overview of the mapping functions of the CycleGAN framework. (A) shows the mapping of a low-density mammogram (source domain) via generator G to a high-density mammogram (target domain). (B) depicts the mapping of a synthetic high-density mammogram back to the low-density source domain via generator F .

discriminator, binary cross entropy is used to compute the adversarial loss as shown below, which is back-propagated to the respective generator network.

$$\begin{aligned}\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) \\ = \mathbb{E}_{y \sim p_{\text{data}}(y)} [\log D_Y(y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log (1 - D_Y(G(x)))]\end{aligned}$$

CycleGAN further contains two cycle-consistency losses defined as $L1$ reconstruction loss between (i) the source image x and the reconstructed source image $F(G(x)) = \hat{x}$ and between (ii) the target image y and the reconstructed target image $G(F(y)) = \hat{y}$.

$$\begin{aligned}\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] \\ + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1]\end{aligned}$$

As defined by (35), the full loss function of our CycleGAN reads as follows with the λ parameter ($\lambda=10$) weighting the relative importance between cycle-consistency and adversarial losses.

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) + \lambda \mathcal{L}_{\text{cyc}}(G, F)$$

Only healthy mammograms (Normal) were used to train the different CycleGAN models. The main reason was to avoid feature hallucinations that have been shown to occur in cancer imaging when training on images where tumors were present (17, 36).

A total of five CycleGAN models from the three different datasets –BCDR, CSAW and OPTIMAM– were trained (see Figure 3). For the OPTIMAM and the CSAW dataset, a different model was trained for each view (CC and MLO). That was because the anatomic features of CC and MLO are different and more specialized CycleGAN models – focusing only on one view – should learn better translations. However, the small sample size in BCDR dataset made it unfeasible to split the models by view and, for this dataset, a single model was trained combining both CC and MLO views (Figure 3A).

The models were trained using a single GPU (24GB NVIDIA GeForce RTX 3090) for a maximum of 200 epochs, using a batch size of 1 and adjusting the learning rate following the recommendations implemented in the CycleGAN Pytorch framework¹. The models are available inside the mediGAN library (37).

2.3 Mass detection using high-density synthetic data augmentation

2.3.1 Effectiveness of the data augmentation strategies

Data scarcity is an important topic in the cancer imaging field and can substantially impact the performance of AI models. In this regard, the effectiveness of the proposed data augmentation technique will depend on the available training images. Ideally, the mass detection models should be trained with large databases with enough representation of the different breast density categories. However, as described in Table 1, there is a considerable imbalance among BI-RADS categories in the datasets. To evaluate the effectiveness of the proposed data augmentation under different data availability conditions, we have designed four different training scenarios.

First, we analyzed the impact of the dataset size. Both a large and a small public dataset were selected to train the detection models, respectively, in high- and low-data regime scenarios. The OPTIMAM Hologic database was selected to investigate the first scenario, comprising more than 3000 mammograms with annotated masses. The low-regime scenario was simulated using the INbreast dataset, which contains 107 FFDMs with annotated masses.

Second, we evaluated how well the synthetic images were able to simulate extremely dense breasts (BI-RADS D) during

¹ <https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>

training. To this end, we trained the detection models twice. First, removing the real BI-RADS D images from the training set and keeping only real mammograms from BI-RADS A, B, and C. In this training setup, the model did not see any real BI-RADS D images during training. Second, including the real BI-RADS D mammograms in the training. In this last scenario, the OPTIMAM Hologic detection models used 25% of the real BI-RADS D mammograms available for training while the remaining ones were used for testing and validation purposes. Note that the INbreast dataset has only 8 real BI-RADS D mammograms available for training.

2.3.2 Training and testing

Our baseline model for mass detection in FFDM is a Deformable DETR (38) with a ResNet50 (39) feature extraction backbone. The model choice was based on the good performance of Deformable DETR in our previous comparative study (34).

First, the baseline model was trained without synthetic data augmentation. Then, four mass detection models were trained with different data augmentation strategies as follows. Three mass detection models, named *BC-Aug*, *CS-Aug* and *OP-Aug*, used synthetic images from a single CycleGAN model with a proportion of 1:1 per mammogram – 1 real and 1 synthetic. The fourth detection model, named *OP-CS-BC-Aug*, included a combination of synthetic images from all the CycleGAN models with a proportion of 1:3 per mammogram – 1 real and 3 synthetic. Thus, the combined detection model was trained with a proportion of synthetic data three times higher than the other three models.

The *BC-Aug* detection model was trained using synthetic data generated from the *BC-All* CycleGAN (Figure 3). Similarly, the *OP-Aug* and the *CS-Aug* detection models were trained with synthetic data from the corresponding CycleGAN models. Note that *BC-All* generates both the MLO and CC views, whereas the CSAW and OPTIMAM have two independent models for each view: *CS-CC*, *CS-MLO*, *OP-CC* and *OP-MLO*.

Only random flipping was used as additional data augmentation. All models were trained five times, that is, using five different random seeds, and evaluated by averaging the results across seeds. A single GPU was used (24GB NVIDIA GeForce RTX 3090) to train each model for a maximum of 60 epochs, using a batch size of 1 and adjusting the learning rate following the implementation recommendations of the MMDetection framework².

2.3.3 Evaluation metrics and statistical significance tests

The area under the curve (AUC) of the Free-response Receiver Operating Characteristic (FROC) curve (40) was used

to compare the baseline with the different data augmentation strategies. The AUC was computed by varying the confidence threshold of each bounding box in a range of FPPI $\in (0, 1)$ (False Positives per Image). If the Intersection-over-Union (IoU) of the prediction and the ground truth was greater than 10%, then the bounding box was considered as a True Positive (TP) (41).

To assess statistical differences in AUCs between the baseline and the models trained with different data augmentation strategies, we used the paired version of DeLong's test for ROC curves (42). To do so, we defined a maximum of 10 False Positives per Image (FPPI) and compared the detection scores of the baseline with the augmented models. The statistical analysis was performed using the code from the fast implementation of DeLong by (43).

2.4 Reader study

A reader study involving two breast radiologists and one surgical oncologist specializing in breast disease was conducted to determine whether the synthetic images were distinguishable from the real ones as a proxy for perceptual realism. The readers were two breast radiologists from different hospitals in Spain, with +9 (Reader A) and +7 (Reader B) years of experience and the surgical oncologist from a hospital in Poland with +12 years of experience in image guided breast biopsy and lesion localization techniques (Reader C).

The reader study contained 180 high-density mammograms balanced by view and dataset. A total of 90 images were original BI-RADS D mammograms: 30 from OPTIMAM Hologic, 30 from CSAW and 30 from BCDR dataset. The other 90 images were synthetic mammograms generated with the different CycleGAN models: 30 images from *OP-CC* and *OP-MLO* models, 30 from *CS-CC* and *CS-MLO* models, and 30 from *BC-All* model. The original low-density mammograms used to generate the synthetic images were randomly selected from BI-RADS A OPTIMAM Hologic dataset.

The reader study was designed as a stand-alone ImageJ³ plugin. A single mammogram was displayed at a time (Figure 4B) next to a multiple-choice panel (Figure 4A) to assign a label based on the certainty of the image being synthetic (Fake) or original (Real). The 6 different choices were converted to equally-distributed probabilities of (0.95, 0.77, 0.59, 0.41, 0.23, 0.05) to compute the ROC curve of each reader as in Alyafi et al. (44). No feedback was given to the readers during the assessment to avoid the identification of artifacts of synthetic images.

The images were resized to a maximum 532 pixels height to avoid the identification of the checkerboard artifacts and the lack of sharpness of synthetic mammograms, which is related to

² <https://github.com/open-mmlab/mmdetection>

³ <https://imagej.nih.gov/ij/>

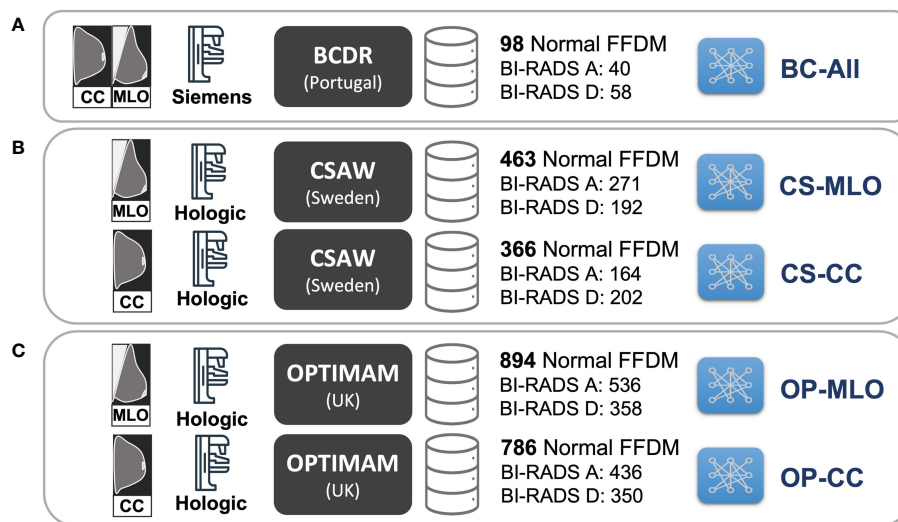


FIGURE 3

Training setup for the different CycleGAN models. (A) *BC-All* model for both CC and MLO views trained with 98 normal FFDMs from BCDR dataset. (B) two models for CC (*CS-CC*) and MLO (*CS-MLO*) views trained with 463 and 366 normal FFDMs from CSAW dataset. (C) two models for CC (*OP-CC*) and MLO (*OP-MLO*) views trained with 894 and 786 normal FFDMs from OPTIMAM Hologic dataset.

upsampling (45) in GANs. The goal of the reader test was to evaluate the anatomically-plausible realism of the synthetic images rather than the noise and common artifacts of the CycleGAN models. Additionally, we asked radiologists to identify the artifacts and common pitfalls of the synthetic images after performing the reader study (Section 4.3.2.1).

3 Results

3.1 Evaluation for the CycleGAN models

The Fréchet Inception Distance (FID) (46) is a useful metric to measure the quality of the synthetic mammograms and compare the synthetic models with each other. The FID was calculated between two different sets of images⁴. Since FID is not an absolute measure, we defined lower and upper bounds using real mammograms. First, the lower bound was defined as the FID between two different splits of real BI-RADS D mammograms from the same dataset. Second, the upper bound was given by the FID between real BI-RADS A and BI-RADS D mammograms. The synthetic BI-RADS D images in our evaluation set were generated from real BI-RADS A mammograms using the five different CycleGAN models (*BC-*

All, *CS-CC*, *CS-MLO*, *OP-CC* and *OP-MLO*). The BI-RADS A mammograms were from the OPTIMAM Hologic and INbreast datasets, the training datasets for the mass detection models. In such manner, we would like to evaluate if the CycleGAN models were able to translate from the source domains – OPTIMAM Hologic and INbreast – to the target domains – OPTIMAM Hologic, CSAW or BCDR – with enough fidelity.

In Table 2, the FID between different groups of images is shown. Overall, the FID score was lower for the synthetic CC mammograms than for the MLO view. Ideally, the FID between synthetic and real BI-RADS D mammograms should be as close as possible to the lower bound and do not exceed the upper bound. As an example, the FID scores in CSAW CC view between real and synthetic BI-RADS D for both OPTIMAM Hologic and INbreast input images, were 99.95 and 124.84, which are between the lower bound (42.57) and the upper bound (142.34). For the synthetic BI-RADS D images generated from OPTIMAM Hologic, both in OPTIMAM Hologic and CSAW CycleGANs, the FID between the real and the synthetic BI-RADS D mammograms lies between the FID bounds. Only in BCDR, the FID score was greater than the FID between both original images.

On the contrary, for the synthetic BI-RADS D images generated from INbreast, the FID was lower in BCDR than in OPTIMAM Hologic and CSAW. Both BCDR and INbreast were acquired with Siemens scanners, while OPTIMAM and CSAW were acquired with Hologic Inc. scanners. This indicates a less pronounced domain difference between BCDR and INbreast, which aligns with the correspondingly smaller FID. Lastly, the

⁴ GitHub repository used to compute FID: <https://github.com/mseitzer/pytorch-fid> (commit 3d604a2)

synthetic images from the OPTIMAM Hologic CycleGAN models lie outside the bounds when the input images came from INbreast. In section 4.2, we will evaluate the impact on mass detection performance when synthetic images with high FID scores are used as data augmentation.

3.1.1 Qualitative analysis of the synthetic images

In Figure 5 there are some sample high-density mammograms generated with the different CycleGAN models. The first row (Figure 5A) corresponds to the CC view and the second (Figure 5B) to MLO. The first column is the original BI-RADS A FFDMs from OPTIMAM Hologic scanner. The next columns are the synthetic FFDMs from the different CycleGAN models. By visual inspection, one can see that the synthetic images did not remove the masses in the original mammograms, which enabled their usage for data augmentation in the mass detection training. We can also confirm that the synthetic images are able to properly translate from source to target

domain. The low-to-high-density image translation was applied for the input BI-RADS A, B, or C. In higher density mammograms, such as BI-RADS C, the changes were more subtle and less density was added to the output image.

3.2 Mass detection performance

All models were evaluated in independent sets of 120 mammograms of each BI-RADS category from the OPTIMAM Hologic dataset. As our objective was to improve the detection performance in BI-RADS D mammograms, we focused on the performance gain in BI-RADS D mammograms. The evaluation metrics for the other BI-RADS categories (A, B and C) can be found in the [Supplementary Material](#).

3.2.1 Large data availability scenario

The corresponding FROC curves of the detection models trained with OPTIMAM Hologic are shown in Figure 6. All

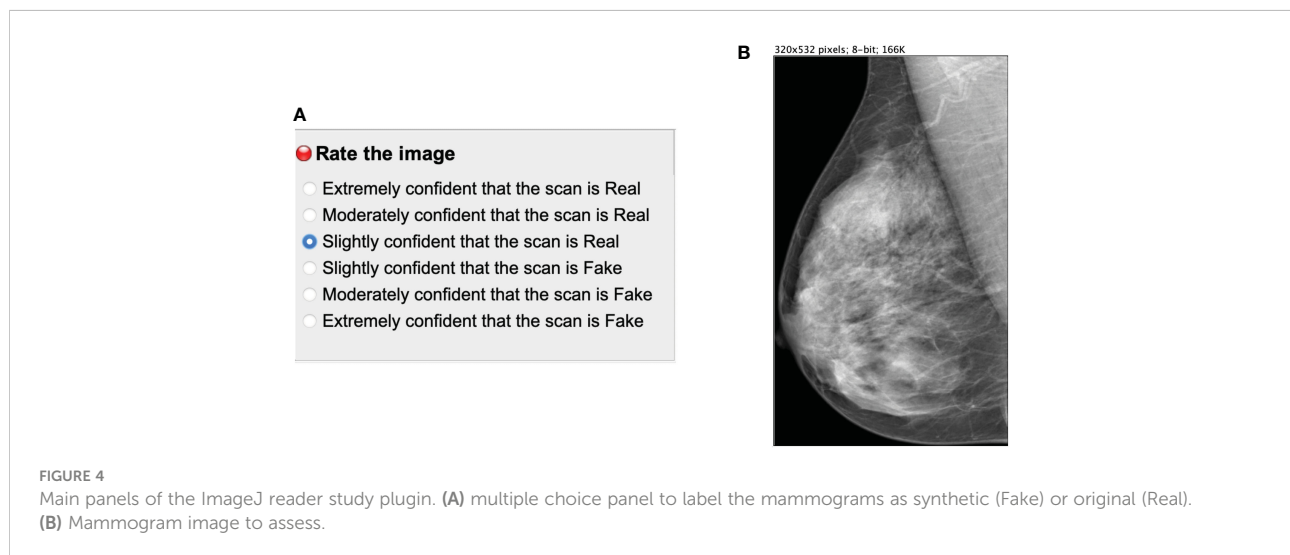


FIGURE 4
Main panels of the ImageJ reader study plugin. (A) multiple choice panel to label the mammograms as synthetic (Fake) or original (Real). (B) Mammogram image to assess.

TABLE 2 Fréchet Inception Distance between different sets of images.

	CycleGAN Dataset	View	Real BI-RADS D (lower bound)	Synthetic BI-RADS D (input BI-RADS A from OPTIMAM Hologic)	Synthetic BI-RADS D (input BI-RADS A from INbreast)	Real BI-RADS A (upper bound)
Real BI-RADS D	BCDR	CC + MLO	66.16	149.10 (BC-All)	103.98 (BC-All)	142.61
		CC	42.57	99.95 (CS-CC)	124.84 (CS-CC)	142.34
	CSAW	MLO	73.54	165.48 (CS-MLO)	183.89 (CS-MLO)	206.04
		CC	34.17	73.16 (OP-CC)	132.04 (OP-CC)	107.99
	OPTIMAM Hologic	MLO	57.24	109.68 (OP-MLO)	175.63 (OP-MLO)	140.93
		CC				

The lower bound was defined as the FID between two different splits of real BI-RADS D mammograms from the same CycleGAN dataset. Similarly, the upper bound was given by the FID between real BI-RADS A and BI-RADS D mammograms. The synthetic BI-RADS D images were generated from real BI-RADS A mammograms from OPTIMAM Hologic and INbreast datasets. The different CycleGAN models (BC-All, CS-CC, CS-MLO, OP-CC and OP-MLO) were used to generate the synthetic images. Bold values indicates the FID values from the synthetic images.

detection models were evaluated on BI-RADS D test set from OPTIMAM Hologic while the INbreast dataset was used as an external validation set.

The models trained using only synthetic BI-RADS D mammograms in training obtained more benefit from the high-density data augmentation. [Table 3](#) summarizes the gains in AUC and the *p-values* for the best performing data augmentation strategies. When only synthetic BI-RADS D images were present in training, the combined data augmentation strategy (*OP-CS-BC-Aug*) obtained a gain of +1.24, increasing the AUC from 79.71% to 80.95% with a *p-value* of 0.0696. The *OP-CS-BC-Aug* obtained a +2.95 gain in AUC in the external validation dataset – INbreast. This confirmed that the resulting model is more robust in the presence of domain-shifts compared to the baseline. However, the best performing data augmentation strategy for the INbreast test set was the *BC-Aug*, with a total gain of +4.15 in AUC. The models trained with real and synthetic BI-RADS D mammograms obtained less gain in the OPTIMAM Hologic test set. To this end, the most substantial gain was achieved with the *BC-Aug* model, with +0.5 in AUC and a *p-value* of 0.2277, which indicates it is not statistically significant. In the external test set, the model with the highest increase in AUC was the *OP-*

Aug model, with a gain of +1.45 in INbreast. This increase in AUC in the external test set showed that the models benefited from high-density synthetic data augmentation even when the training data is large and contains real BI-RADS D mammograms. The detailed metrics for the other data augmentation strategies can be found in [Supplementary Material \(Table D\)](#).

3.2.2 Low data availability scenario

The corresponding FROC curves of the detection models trained on the INbreast dataset are shown in [Figure 7](#). All mass detection models were evaluated on the BI-RADS D test set from OPTIMAM Hologic. [Table 4](#) summarizes the gains in AUC and the *p-values* for the best performing data augmentation strategies in INbreast. The detailed metrics for the other data augmentation strategies can be found in [Supplementary Material \(Table E\)](#).

The models trained using only synthetic BI-RADS D mammograms in training improved their performance in two out of four data augmentation strategies, namely, *OP-Aug* and *OP-CS-BC-Aug*. The *OP-Aug* model increased the AUC from 42.59% to 45.41%, a gain of +2.81 with a *p-value* smaller than 0.05 with respect to the baseline model. The *OP-CS-BC-Aug*

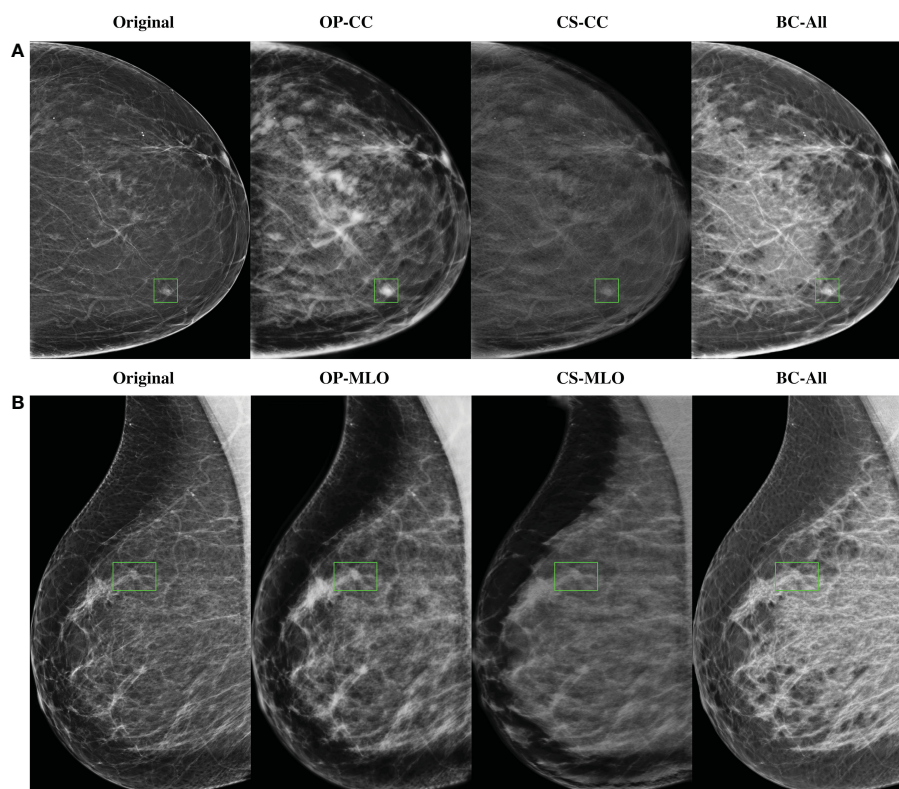
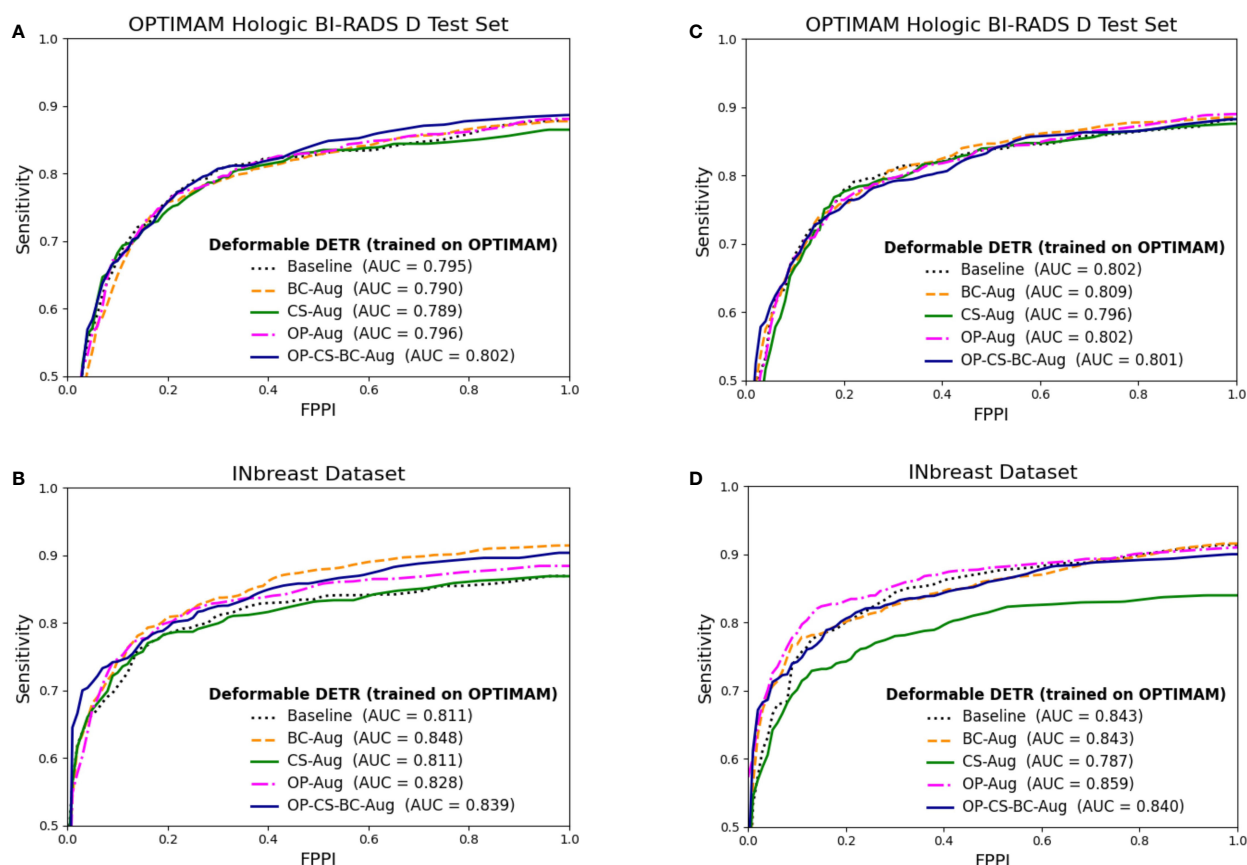


FIGURE 5

Samples of high-density synthetic mammograms generated with the different CycleGAN models. On the left, the input BI-RADS A mammogram from OPTIMAM Hologic dataset. (A) CC view, (B) MLO view.



Only synthetic BI-RADS D in training

Real and synthetic BI-RADS D in training

FIGURE 6

FROC curves for the mass detection models trained using OPTIMAM Hologic database. The baseline model did not include synthetic data augmentation in the training. The *BC-Aug*, *CS-Aug*, *OP-Aug* and the *OP-CS-BC-Aug* models used synthetic images from the corresponding CycleGAN models. (A, C) The OPTIMAM Hologic test set contained 120 BI-RADS D mammograms with annotated masses, (B, D) The INbreast test set containing 107 mammograms with annotated masses from different breast densities, used for external validation. (C, D) The models were trained with 41 additional real BI-RADS D mammograms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

model increased the baseline AUC by +1.05. When training the models with real and synthetic BI-RADS D mammograms, only the *OP-Aug* model obtained a gain in performance with respect to the baseline. However, the gain is higher than in other scenarios, increasing the AUC from 44.59% to 48.84%, a gain of +4.25.

The AUC values in the OPTIMAM Hologic BI-RADS D test set are low in comparison with the AUCs in the large availability scenario. A first reason is that the INbreast models were trained with only 116 mammograms with masses. Second, the test set is from a different domain (OPTIMAM Hologic) due to the fact that the entire INbreast dataset was used as training set. The latter was motivated by the fact that the number of BI-RADS D mammograms available in the INbreast dataset (Table 1) was deemed insufficient for representative performance evaluation.

3.3 Reader study outcomes

3.3.1 Receiver operating characteristic curves

The AUC of each reader ROC curve can be found in Table 5. On average, the synthetic mammograms of CC view from OPTIMAM CycleGAN were the most difficult to discriminate from original mammograms (0.615 AUC). Overall, the CC view looked more realistic to all readers than the MLO view. The BCDR model was the easiest for recognition of synthetic images, with an AUC of 0.824 in CC view and 0.954 in MLO view.

3.3.2 Qualitative analysis of synthetic FFDM

After completing the study, the readers evaluated the realism of the synthetic images and identified the common artifacts and failures.

TABLE 3 Performance values and statistical significance test results of the best data augmentation strategies for the models trained with OPTIMAM Hologic database.

		Only synthetic BI-RADS D in training			Synthetic and real BI-RADS D in training			
		FROC AUC	Gain	<i>p-value</i>	FROC AUC	Gain	<i>p-value</i>	
OPTIMAM Hologic BI-RADS D Test Set	Baseline	79.71% (78.44, 80.98)	Ref	Ref	80.60% (79.20, 82.00)	Ref	Ref	
	BC-Aug	79.62% (77.83, 81.41)	-0.09	0.0064	81.10% (80.40, 81.80)	+0.50	0.2277	0.2277
	OP-Aug	79.86% (78.30, 81.42)	+0.15	0.8269	80.75% (78.77, 82.73)	+0.15	0.5599	0.5599
	OP-CS-BC-Aug	80.95% (79.63, 82.27)	+1.24	0.0696	80.76% (79.92, 81.60)	+0.16	0.7921	0.7921
INbreast Dataset (external validation)	Baseline	81.51% (78.93, 84.09)	Ref	Ref	84.71% (83.39, 86.03)	Ref	Ref	
	BC-Aug	85.66% (81.91, 89.41)	+4.15	0.0002	84.88% (82.86, 86.90)	+0.17	0.1666	0.1666
	OP-Aug	83.45% (80.03, 86.87)	+1.94	6.08e-05	86.16% (83.37, 88.95)	+1.45	0.0041	0.0041
	OP-CS-BC-Aug	84.47% (82.32, 86.62)	+2.95	0.0008	84.29% (82.22, 86.36)	-0.42	0.0162	0.0162

The columns on the left correspond to the models trained without real BI-RADS D mammograms. The baseline models were trained without synthetic images. The 95% Confidence Intervals of the FROC AUC are in parenthesis. The *p-value* was computed using the DeLong method with a maximum of 10 FPPI. Bold values correspond to the best performing strategy. Ref corresponds to the reference method.

3.3.2.1 Common artifacts and failures

The most unrealistic features in order of importance were: 1) big concentrations of glandular tissue adjacent to pectoral muscle (Figure 8D); 2) dark small dots or ovals in the image like perforations (Figure 8A); 3) linear fragmentation of muscles; 4) distorted nipples (Figure 8C); and 5) lack of glandular tissue behind the nipple (Figure 8B). Some synthetic images missed to preserve the prepectoral fat or smooth contours at the interface with subcutaneous fat.

3.3.2.2 Realistic mammograms

Some synthetic high-density FFDMs that looked realistic to the radiologist are shown in Figure 9. The features that improved realism of synthetic images were 1) the presence of linear microcalcifications or roundish calcifications; 2) lymph nodes in the correct areas; 3) post-biopsy tissue markers; and 4) the correct distribution of dense tissue.

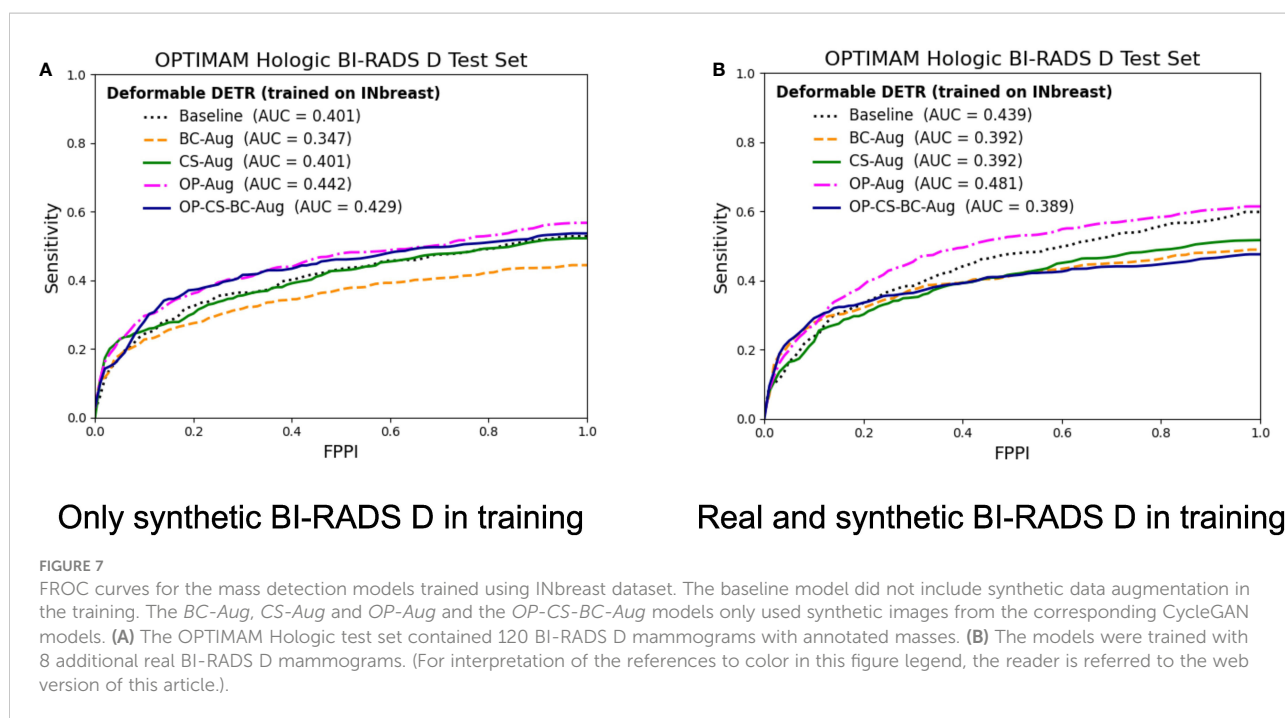
4 Discussion

The contributions in this study are two fold. First, synthesize high-resolution high-density FFDMs from different domains. Second, use the synthetic images as data augmentation to

improve the mass detection performance in high-density breasts and multi-center datasets.

In our image synthesis experiments, we trained a total of five CycleGAN models using FFDMs from three different datasets to perform low-to-high density translation from real mammograms. Only healthy mammograms were used to train the image-to-image translation models. By training with healthy images, the models learned the healthy data distribution, hence, minimizing any risk of inserting hallucinated lesions into the synthetic images. As our goal was to use the synthetic mammograms to improve the mass detection, we confirmed that the CycleGAN models did not remove the masses from the original mammograms (see Figure 5).

To assess the quality of the synthetic images we calculated the commonly used FID metric between synthetic and real high-density mammograms. Since FID is not an absolute measure, we defined lower and upper bounds using real mammograms. The closer the synthetic images are to the real BI-RADS D images, the closer the FID should be to the lower bound. After evaluating the FID metric of the different CycleGAN models, we observed that the FID for the CC view was better than the one of the MLO view. The difficulty of synthesizing MLO view was previously mentioned in Korkinof et al. (24), most probably because MLO has greater complexity and more anatomical information than the CC view. However, the *OP-MLO* and *CS-MLO* models still



had an FID score between the lower and upper bounds when the input images were from OPTIMAM Hologic.

The *BC-All* CycleGAN had the largest FID score when the input images were from OPTIMAM Hologic, and the smallest FID when the input images were from INbreast. This is in line with the scanner manufacturer differences between both datasets. BCDR and INbreast were acquired with Siemens scanners. On the other hand, OPTIMAM and CSAW were acquired with Hologic Inc. scanners. Moreover, BCDR contains old digitized film mammograms while OPTIMAM is a Hologic digital mammography dataset with very specific image characteristics, i.e., it is very sharp, usually shows lymph nodes very well and it does visualize some skin of the breast.

To further evaluate the clinical realism of the synthetic images, we performed a reader study involving two breast

radiologists and one surgical oncologist. When the CycleGAN models trained using OPTIMAM Hologic and CSAW datasets (*OP-CC*, *OP-MLO*, *CS-CC*, *CS-MLO*) were used to insert density onto low density mammograms from OPTIMAM Hologic, it was much more difficult for the readers to differentiate between original and synthetic mammograms. In that case, the readers had to look for anatomical disparities and inadequacies to spot the difference. Both OPTIMAM and CSAW FFDMs were acquired with an Hologic scanner. On the other hand, all the readers could easily identify the synthetic images generated with the *BC-All* model. As previously mentioned, BCDR contains old digitized film mammograms acquired with a Siemens scanner. Considering that, we can conclude that domain disparities between the acquisition settings of the source and the target domains have a big impact on the perceptual realism of the

TABLE 4 Performance values and statistical significance of the best data augmentation strategies for the models trained with INbreast dataset.

		Only synthetic BI-RADS D in training			Synthetic and real BI-RADS D in training		
		FROC AUC	Gain	<i>p</i> -value	FROC AUC	Gain	<i>p</i> -value
OPTIMAM Hologic BI-RADS D Test Set (external validation)	Baseline	42.59% (39.73, 45.45)	Ref	Ref	44.59% (42.87, 46.31)	Ref	Ref
	OP-Aug	45.41% (42.70, 48.12)	+2.81	4.58e-22	48.84% (46.21, 51.47)	+4.25	1.43e-27
	OP-CS-BC-Aug	43.65% (40.64, 46.66)	+1.05	2.32e-11	39.74% (35.72, 43.76)	-4.85	9.40e-10

The columns on the left correspond to the models trained without real BI-RADS D mammograms. The baseline models were trained without synthetic images. The 95% Confidence Intervals of the FROC AUC are in parenthesis. The *p*-value was computed using the DeLong's method with a maximum of 10 FPPI. Bold values correspond to the best performing strategy. Ref corresponds to the reference method.

TABLE 5 Reader test results: Area Under the Curve (AUC) from the Receiver Operating Characteristics (ROC) curve of each CycleGAN model and view (CC, MLO).

	OPTIMAM		CSAW		BCDR	
	CC	MLO	CC	MLO	CC	MLO
Reader A	0.580	0.664	0.651	0.718	0.818	0.962
Reader B	0.576	0.893	0.887	0.758	0.871	0.956
Reader C	0.689	0.673	0.489	0.636	0.784	0.944
Average ± std	0.615 ± 0.052	0.743 ± 0.105	0.675 ± 0.163	0.704 ± 0.050	0.824 ± 0.037	0.954 ± 0.007

Reader A: 9+ years of experience as a breast radiologist. Reader B: 7+ years of experience as a breast radiologist. Reader C: surgical oncologist with +12 years of experience in image guided breast biopsy and lesion localization techniques.

synthetic images. In addition, the synthetic mammograms from the MLO view were easier to identify than the ones from the CC view and this observation correlates with the higher FID scores of the synthetic images from the MLO CycleGAN models (*OP-MLO* and *CS-MLO*).

In our mass detection experiments, different data augmentation strategies were tested to improve the mass detection of the baseline model. First, three mass detection models were trained using synthetic images from the different

CycleGANs with a proportion of 1:1 – 1 real and 1 synthetic. Second, a fourth mass detection model was trained with synthetic images from all the CycleGAN models, the *OP-CS-BC-Aug*, with a proportion of 1:3 – 1 real and 3 synthetic. To evaluate the effect of the different data augmentation strategies under diverse training conditions, we defined four training scenarios for the mass detection models. Two scenarios involving the amount of data available for training, and the other two involving the exclusion or inclusion of the real BI-

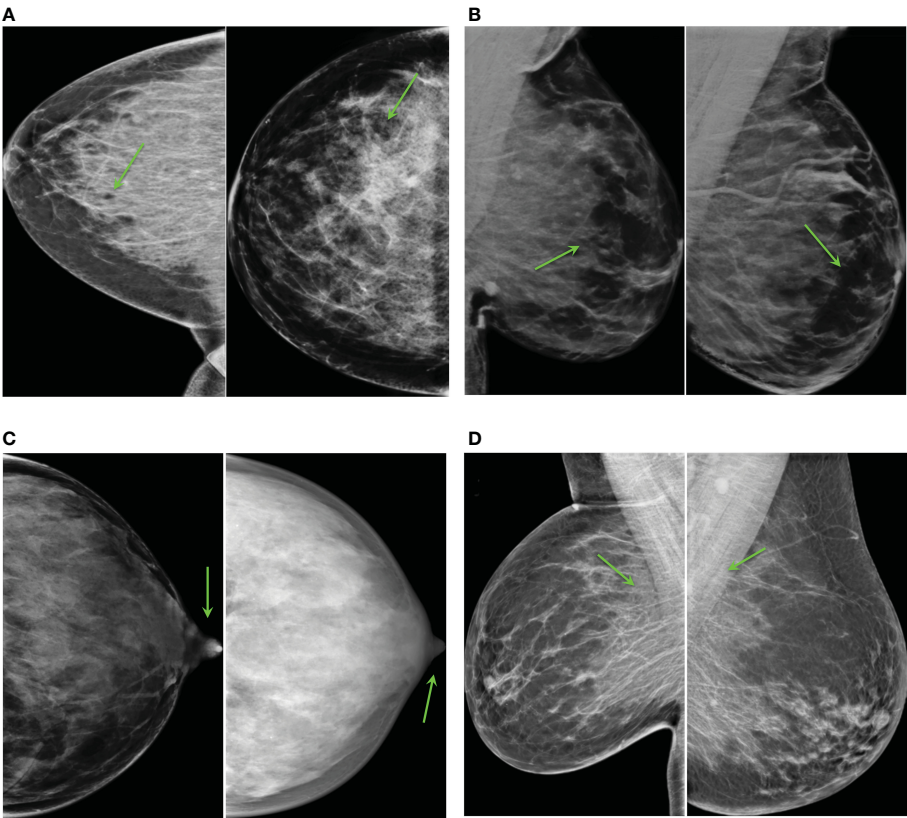


FIGURE 8
Common artifacts of synthetic high-density mammograms, **(A)** dark dots or ovals, **(B)** lack of glandular tissue behind the nipple, **(C)** distorted nipples, **(D)** big concentrations of glandular tissue adjacent to the pectoral muscle.

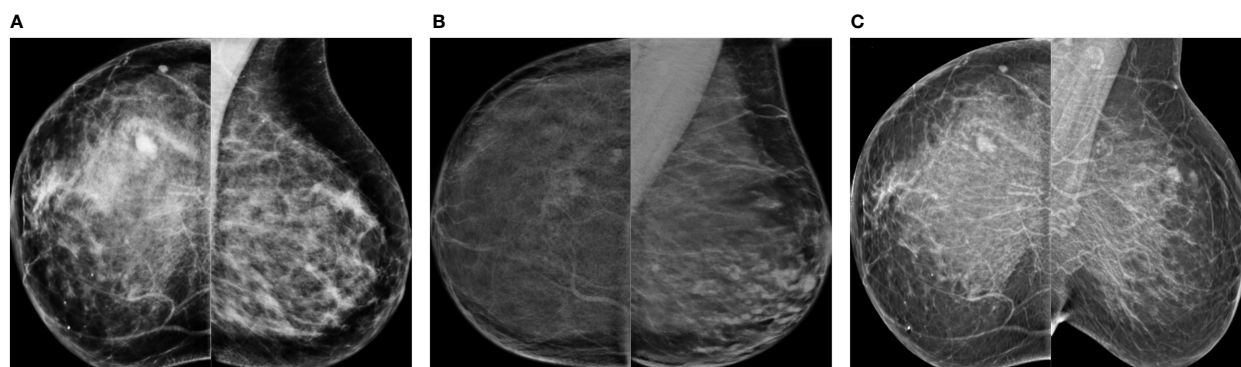


FIGURE 9

Synthetic high-density mammograms that looked realistic to the radiologist. Samples from the different CycleGAN models: (A) *OP-CC* and *OP-MLO*, (B) *CS-CC* and *CS-MLO*, (C) *BC-All*.

RADS D mammograms in training. By excluding the real BI-RADS D mammograms, we tested if the synthetic high-density mammograms could replace the real BI-RADS D images, simulating a plausible clinical scenario in which no BI-RADS D images were available for training. In the large availability scenario, the data augmentation strategies did not improve the baseline performance with statistically significant difference. However, the synthetic data helped the models to generalize better and increased the performance in the external dataset (INbreast). When training only with synthetic BI-RADS D images, the *BC-Aug* model improved the AUC by 4.15% in INbreast. When the real BI-RADS D images were present in training, the *OP-Aug* model improved the AUC by 1.45%. In the low data availability scenario, the *OP-Aug* data augmentation strategy improved the AUC significantly for both scenarios in the out-of-domain test set.

As we hypothesized, the CycleGAN models not only learned how to translate from low-to-high density but also preserved the domain characteristics from their respective training datasets during translation. The domain characteristics comprise the differences in image quality, acquisition settings and scanner manufacturers.

Based on our experimental results, we did not observe a consistent association between the FID scores and the success of the corresponding data augmentation in mass detection. For instance, in the low-data availability scenario, the CycleGANs trained on OPTIMAM Hologic had a comparably high FID but the *OP-Aug* detection model obtained the highest increase in the AUC. Moreover, the *OP-Aug* detection model yielded a consistent improvement in the four training scenarios. The *OP-CC* and *OP-MLO* CycleGANs were trained with a larger number of training images than the CSAW and BCDR CycleGANs. Overall, the number of images used to train the CycleGAN models seemed to have a higher importance for the detection performance gain compared to the FID score.

One of the limitations of our work involves the availability of healthy BI-RADS D mammograms to train the generative models. From the results, the data augmentation that obtained more consistent gains under all four scenarios was the *OP-Aug*, using CycleGANs trained with the OPTIMAM Hologic dataset. As shown in Figure 3, *OP-MLO* and *OP-CC* were trained with more images than the other CycleGANs, showing the big impact of data for training more robust CycleGANs. In this regard, fairer comparisons among generative models could be achieved if more images from the CSAW and BCDR datasets would be available.

Data scarcity is one of the most common and general limitations in medical imaging AI research. In tasks where patch-based approaches are sufficient, extracting multiple samples from one scan can aid to overcome the data scarcity issue (47). However, this limitation is further exacerbated in tasks where a single subject scan can only be used as a single sample in the training of deep learning methods. Moreover, the high resolution nature of the breast FFDMs makes it even more challenging. We believe that this study is not only an important step towards mitigating data scarcity and class imbalance, but also demonstrates the importance of fair AI in clinical practice. However, there is still room for improvement to increase the fairness of AI models for women with high-density breasts in mammography screening.

Future research may focus on evaluating the potential of the high-density synthetic mammograms in other downstream tasks. We analyzed the high-density mammogram synthesis *via* the downstream task of breast mass detection. However, there is a multitude of further applications of our proposed low-to-high breast density translation. For instance, our method can be applied to breast mass segmentation or tumor malignancy classification. Furthermore, our synthetic images can expand the radiologist curricular and professional training programs. For instance, for training purposes radiologist candidates would

need to detect and accurately annotate the same lesion in both high and low-density breast, which allows to measure inter- and intra-observer variability per lesion and breast density level. This way our method provides flexibility to personalize and adjust the radiologist training to specific scenarios (e.g. very high density breasts with very small-sized tumor lesions).

5 Conclusion

In this study, we evaluated different CycleGAN models for high-density FFDM synthesis from three different datasets and acquisition pipelines, comprising two scanner manufacturers – i.e. Hologic and Siemens. Moreover, we applied different synthetic data augmentation strategies to improve the mass detection performance of a deep-learning based model. Even though the improvements were not always statistically significant for models trained in the large data availability scenario, the results demonstrated that the data augmentation helped to improve the mass detection in out-of-domain datasets, improving the domain generalization of the final model. The models trained in the low data availability scenario obtained more benefit from the data augmentation, with a maximum gain of 4.25% in AUC for the model trained with synthetic images generated with the OPTIMAM Hologic CycleGANs. Finally, a reader study involving three expert radiologists evaluated the perceptual realism of the synthetic mammograms, concluding that the quality of CC view synthetic images is higher than the mammograms from MLO view. Our study is the first one to synthesize high-resolution FFDMs with increased density and showed the potential of including the generated images in the data augmentation pipeline to improve the generalization and performance of downstream tasks using mammography images, such as mass detection.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: A subset of the OPTIMAM Database was obtained as part of the data-sharing agreement with the University of Barcelona in 2021. Requests to access these datasets should be directed to Cancer Research Horizons, horizons@cancer.org.uk. The CycleGAN models and weights mentioned in this study can be found in the medigan library (<https://github.com/RichardObi/medigan>).

Ethics statement

Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

LG, KK, RO, and OD contributed to the conception and design of the study. LG organized the datasets, performed the experiments and designed the reader study. MB, AC, and JR participated in the reader study and clinical assessment of the results. FS contributed with one dataset (CSAW) in this study. LI and KL supervised the research. All authors contributed to article revision and read and approved the submitted version.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952103. This work was supported in part by the MICINN Grant RTI2018-095232-B-C21 and 2017 SGR 1742.

Acknowledgments

A subset of the OPTIMAM database was obtained as part of the data-sharing agreement with the University of Barcelona in 2021. Special thanks to Dr Melissa Hill (Volpara Health) for sharing the breast density information of the OPTIMAM subset and to Dr Premkumar Elanvogan for sharing the base code to prepare the ImageJ plugin for the reader study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.1044496/full#supplementary-material>

References

- D'Orsi C, Sickles E, Mendelson E, Morris E. *ACR BI-RADS® Atlas, breast imaging reporting and data system (ACR BI-RADS® Magnetic Resonance Imaging)*. Reston, VA: American College of Radiology (2013).
- Pesce K, Tajerian M, Chico M, Swiecicki M, Boietti B, Frangella M, et al. Interobserver and intraobserver variability in determining breast density according to the fifth edition of the BI-RADS® Atlas. *Radiology* (2020) 2020:481–6. doi: 10.1016/j.rxeng.2020.04.004
- Sprague BL, Gangnon RE, Burt V, Trentham-Dietz A, Hampton JM, Wellman RD, et al. Prevalence of mammographically dense breasts in the united states. *JNCI: J Natl Cancer Institute* (2014) 106(10). doi: 10.1093/jnci/dju255
- Han Y, Moore JX, Colditz GA, Toriola AT. Family history of breast cancer and mammographic breast density in premenopausal women. *JAMA Network Open* (2022) 5(2):e2148983–e2148983. doi: 10.1001/jamanetworkopen.2021.48983
- Kerlikowske K, Zhu W, Tosteson AN, Sprague BL, Tice JA, Lehman CD, et al. Identifying women with dense breasts at high risk for interval cancer: A cohort study. *Ann Internal Med* (2015) 162(10):673–81. doi: 10.7326/M14-1465
- Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *New Engl J Med* (2007) 356:227–36. doi: 10.1056/NEJMoa062790
- Strand F, Azavedo E, Hellgren R, Humphreys K, Eriksson M, Shepherd J, et al. Localized mammographic density is associated with interval cancer and large breast cancer: A nested case-control study. *Breast Cancer Res* (2019) 21:1–9. doi: 10.1186/s13058-019-1099-y
- Seely JM, Peddle SE, Yang H, Chiarelli AM, McCallum M, Narasimhan G, et al. Breast density and risk of interval cancers: The effect of annual versus biennial screening mammography policies in Canada. *Can Assoc Radiologists J* (2022) 73(1):90–100. doi: 10.1177/08465371211027958
- Gordon PB. The impact of dense breasts on the stage of breast cancer at diagnosis: A review and options for supplemental screening. *Curr Oncol* (2022) 29(5):3595–636. doi: 10.3390/curroncol29050291
- McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiol Biomarkers Prev* (2006) 15(6):1159–69. doi: 10.1158/1055-9965.EPI-06-0034
- Chiu SYH, Duffy S, Yen AMF, Tabár L, Smith RA, Chen HH. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. *Cancer Epidemiology Biomarkers Prev* (2010) 19(5):1219–28. doi: 10.1158/1055-9965.EPI-09-1028
- Engmann NJ, Golmakani MK, Miglioretti DL, Sprague BL, Kerlikowske K. Population-attributable risk proportion of clinical risk factors for breast cancer. *JAMA Oncol* (2017) 3(9):1228–36. doi: 10.1001/jamaoncol.2016.6326
- van der Waal D, Rippling TM, Verbeek AL, Broeders MJ. Breast cancer screening effect across breast density strata: A case-control study. *Int J Cancer* (2016) 2017:41–9. doi: 10.1002/ijc.30430
- Weigel S, Heindel W, Heidrich J, Hense HW, Heidinger O. Digital mammography screening: Sensitivity of the programme dependent on breast density. *Eur Radiol* (2017) 27:2744–51. doi: 10.1007/s00330-016-4636-4
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: *Advances in neural information processing systems*, vol. 27. Montreal, Quebec, Canada: Curran Associates, Inc (2014). p. 2672–80. et al. ZG.
- Oza P, Sharma P, Patel S, Adedoyin F, Bruno A. Image augmentation techniques for mammogram analysis. *J Imaging* (2022) 8(5):141. doi: 10.3390/jimaging8050141
- Osuala R, Kushibar K, Garrucho L, Linardos A, Szafranowska Z, Klein S, et al. Data synthesis and adversarial networks: A review and meta-analysis in cancer imaging. *Medical Image Analysis* (2021) 84. doi: 10.1016/j.media.2022.102704
- AlAmir M, AlGhamdi M. The role of generative adversarial network in medical image analysis: An in-depth survey. *ACM Comput Surv* (2022) 55(5):1–36. doi: 10.1145/3527849
- Becker AS, Jendele L, Skopek O, Berger N, Ghafoor S, Marcon M, et al. Injecting and removing suspicious features in breast imaging with CycleGAN: A pilot study of automated adversarial attacks using neural networks on small images. *Eur J Radiol* (2019), 108649. doi: 10.1016/j.ejrad.2019.108649
- Zakka C, Saheb G, Najem E, Berjawi G. Mammogenesis: Controlled generation of high-resolution mammograms for radiology education. *arXiv (Preprint)* (2020). doi: 10.48550/arXiv.2010.05177
- Desai SD, Giraddi S, Verma N, Gupta P, Ramya S. (2020). Breast cancer detection using gan for limited labeled dataset, in: *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)* Bhimtal, India: IEEE pp. 34–9. doi: 10.1109/CICN49253.2020.9242551
- Jendele L, Skopek O, Becker AS, Konukoglu E. Adversarial augmentation for enhancing classification of mammography images (Internet). *arXiv (Preprint)* (2019). doi: 10.48550/arXiv.1902.07762
- Muramatsu C, Nishio M, Goto T, Oiwa M, Morita T, Yakami M, et al. Improving breast mass classification by shared data with domain transformation using a generative adversarial network. *Comput Biol Med* (2020) 119:103698. doi: 10.1016/j.compbio.2020.103698
- Korkinof D, Rijken T, O'Neill M, Yearsley J, Harvey H, Glocker B. High-resolution mammogram synthesis using progressive generative adversarial networks. *arXiv (Preprint)* (2018). doi: 10.48550/arXiv.1807.03401
- Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of gans for improved quality, stability, and variation. In: *The international conference on learning representations*. Vancouver, BC, Canada: Conference Track Proceedings (2018). doi: 10.48550/arXiv.1710.10196
- Korkinof D, Harvey H, Heindl A, Karpati E, Williams G, Rijken T, et al. Perceived realism of high-resolution generative adversarial network-derived synthetic mammograms. *Radiology. Artif Intell* (2021) 3(2). doi: 10.1148/ryai.2020190181
- Cha KH, Petrick NA, Pezeshk AX, Graff CG, Sharma D, Badal A, et al. Evaluation of data augmentation via synthetic images for improved breast mass detection on mammograms using deep learning. *J Med Imaging* (2019) 7(1):1–9. doi: 10.1117/1.JMI.7.1.012703
- Badano A, Graff CG, Badal A, Sharma D, Zeng R, Samuelson FW, et al. Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial. *JAMA Network Open* (2018) 1(7):e185474–e185474. doi: 10.1001/jamanetworkopen.2018.5474
- Halling-Brown MD, Warren LM, Ward D, Lewis E, Mackenzie A, Wallis MG. OPTIMAM mammography image database: A large-scale resource of mammography images and clinical data. *Radiology: Artif Intell* (2021):e200103. doi: 10.1148/ryai.2020200103
- Dembrower K, Lindholm P, Strand F. A multi-million mammography image dataset and population-based screening cohort for the training and evaluation of deep neural networks—the cohort of screen-aged women (CSAW). *J digital Imaging* (2020) 33:408–13. doi: 10.1007/s10278-019-00278-0
- Keller BM, Chen J, Daye D, Conant EF, Kontos D. Preliminary evaluation of the publicly available laboratory for breast radiodensity assessment (LIBRA) software tool: Comparison of fully automated area and volumetric density measures in a case-control study with digital mammography. *Breast Cancer Res* (2015) 17:1–17. doi: 10.1186/s13058-015-0626-8
- Lopez MG, Posada N, Moura DC, Pollán RR, Valiente JMF, Ortega CS, et al. BCDR: A breast cancer digital repository. In: *15th International Conference on Experimental Mechanics*, Porto/Portugal, 22–27 July 2012 (2012).
- Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS. INbreast: toward a full-field digital mammographic database. *Acad Radiol* (2011) 2012:236–48. doi: 10.1016/j.acra.2011.09.014
- Garrucho L, Kushibar K, Jouide S, Diaz O, Igual L, Lekadir K. Domain generalization in deep learning based mass detection in mammography: A large-scale multi-center study. *Artif Intell Med* (2022) 132:102386. doi: 10.1016/j.artmed.2022.102386
- Zhu J, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *IEEE International conference on computer vision (ICCV)*. Los Alamitos, CA, USA: IEEE Computer Society (2017). p. 2242–51. doi: 10.1109/ICCV.2017.244
- Cohen JP, Luck M, Honari S. Distribution matching losses can hallucinate features in medical image translation. In: *International conference on medical image computing and computer-assisted intervention*. Cham: Springer International Publishing (2018). p. 529–36. doi: 10.1007/978-3-030-00928-1_60
- Osuala R, Skorupko G, Lazrak N, Garrucho L, García E, Joshi S, et al. Medigan: A Python library of pretrained generative models for enriched data access in medical imaging. *J Med Imaging* (2022), arXiv preprint arXiv:2209.14472. doi: 10.48550/arXiv.2209.14472
- Zhu X, Su W, Lu L, Li B, Wang X, Dai J. (2021). Deformable DETR: Deformable transformers for end-to-end object detection, in: *International Conference on Learning Representations*, . doi: 10.48550/arXiv.2010.04159
- He K, Zhang X, Ren S, Sun J. (2016). Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Las Vegas: IEEE). pp. 770–8.
- Bandos AI, Rockette HE, Song T, Gur D. Area under the free-response ROC curve (FROC) and a related summary index. *Biometrics* (2009) 65(1):247–56. doi: 10.1111/j.1541-0420.2008.01049.x
- McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast

cancer screening. *Nature* (2020) 577:89–94. doi: 10.1038/s41586-019-1799-6

42. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* (1988) 44(3):837–45. doi: 10.2307/2531595

43. Sun X, Xu W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Process Letters* (2014) 21(11):1389–139. doi: 10.1109/LSP.2014.2337313

44. Alyafi B, Diaz O, Elangovan P, Vilanova JC, del Riego J, Marti R. Quality analysis of DCGAN-generated mammography lesions. In: Bosmans H, Marshall N, Ongeval CV, editors. *15th International Workshop on Breast Imaging*, vol. 11513. Leuven, Belgium: SPIE (2020). p. 80–5. doi: 10.1117/12.2560473

45. Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts. *Distill* (2016). doi: 10.23915/distill.00003.

46. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In: Guyon I, et al, editors. *Advances in neural information processing systems*, vol. 30. Long Beach, CA, USA - DBLP: Curran Associates, Inc (2017). p. 6629–40.

47. Szafranowska Z, Osuala R, Breier B, Kushibar K, Lekadir K, Diaz O. Sharing generative models instead of private data: a simulation study on mammography patch classification. In: Bosmans H, et al, editors. *16th international workshop on breast imaging (IWBI2022)*, Leuven, Belgium: International Society for Optics and Photonics (SPIE (2022). p. 122860Q. doi: 10.1117/12.2625781



OPEN ACCESS

EDITED BY
Laura Curiel,
University of Calgary,
Canada

REVIEWED BY
Monika Rucinska,
University of Warmia and Mazury in
Olsztyn, Poland
Rony Benson,
Marsleeva Medicty, India

*CORRESPONDENCE
Olexiy I. Aseyev
✉ Olexiy.Aseyev@tbh.net

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 06 December 2022
ACCEPTED 09 January 2023
PUBLISHED 24 January 2023

CITATION
Clark M, Griborio-Guzman AG,
Burute NP, Lubbers S, Anthes ML,
Sadreddini M and Aseyev OI (2023)
When tissue is not the only issue: Poorly
differentiated lung squamous-cell
carcinoma with adrenal, costochondral,
and cardiac metastases – case report.
Front. Oncol. 13:1117024.
doi: 10.3389/fonc.2023.1117024

COPYRIGHT
© 2023 Clark, Griborio-Guzman, Burute,
Lubbers, Anthes, Sadreddini and Aseyev. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

When tissue is not the only issue: Poorly differentiated lung squamous-cell carcinoma with adrenal, costochondral, and cardiac metastases – case report

Megan Clark¹, Andres G. Griborio-Guzman^{1,2},
Nishigandha P. Burute^{1,3}, Sonja Lubbers^{1,4}, Margaret L. Anthes^{1,5,6},
Masoud Sadreddini^{1,2} and Olexiy I. Aseyev^{1,6,7*}

¹Northern Ontario School of Medicine University, Thunder Bay, ON, Canada, ²Division of Cardiology, Department of Internal Medicine, Thunder Bay Regional Health Sciences Centre, Thunder Bay, ON, Canada, ³Department of Diagnostic Imaging, Thunder Bay Regional Health Sciences Centre, Thunder Bay, ON, Canada, ⁴Department of Internal Medicine, Thunder Bay Regional Health Sciences Centre, Thunder Bay, ON, Canada, ⁵Department of Radiation Oncology, Thunder Bay Regional Health Sciences Centre, Thunder Bay, ON, Canada, ⁶Regional Cancer Care Northwest, Thunder Bay, ON, Canada, ⁷Department of Medical Oncology, Thunder Bay Regional Health Sciences Centre, Thunder Bay, ON, Canada

Nonmelanoma skin cancer is the most common cancer in the world, and lung cancer is the leading cause of death from cancer. Histologically, squamous cell carcinoma (SCC) is the second most prevalent type of both skin and lung cancers. We report the case of a 38-year-old female with metastatic, poorly differentiated lung SCC detected on chest X-ray after she presented to the hospital with cough and dyspnea. She had had a 7.5 cm moderately differentiated well-circumscribed posterior scalp SCC completely excised eight years earlier. CT scan showed a large right lung mass, nodular filling defect in the left atrium (LA), and metastases to the adrenal glands and the first rib. Her pulmonary tumor extends to the LA via the right superior pulmonary vein, which is rarely reported in the literature. Ultrasound-guided biopsy of the rib mass showed poorly differentiated SCC. The patient received urgent radiotherapy, given superior vena cava and mainstem bronchus compression. Head CT showed no brain metastasis. A biopsy of the left adrenal initially reported an undifferentiated pleomorphic sarcoma; however, a second pathologist reported it as a poorly differentiated carcinoma of lung origin. At least three pathologists verified the specimen, and it had a PD-L1 test with a 1-49% score. An initial echocardiogram confirmed the LA mass. The patient received a Paclitaxel-Carboplatin-Pembrolizumab regimen as the first-line treatment for metastatic SCC. A repeat echocardiogram after cycle 1 showed a decrease in the size of the tumor in the LA. Almost five months after her initial visit, this young woman's symptoms and performance status have improved post-palliative radiotherapy and chemo-immunotherapy. Follow-up CT showed smaller lung, nodal, adrenal, and costochondral masses, and evidence of necrosis. This case is clinically relevant because it represents a common problem presenting

uncommonly. Moreover, it highlights that ultrasound-guided interventions and medical imaging are essential in directing metastatic cancer diagnosis, treatment, and follow-up, especially when pathology cannot confirm but only presume a specific diagnosis.

KEYWORDS

squamous cell carcinoma (SCC), non-small cell lung cancer (NSCLC), cardiac imaging techniques, undifferentiated pleomorphic sarcoma (UPS), cardiac tumors, ultrasound-guided biopsy, computed tomography (CT) scan, echocardiography

1 Introduction

Our report describes the case of a non-smoker 38-year-old woman with metastatic squamous cell carcinoma (SCC) of the lung and previously diagnosed with cutaneous squamous cell carcinoma (cSCC) at age 30. The patient presented to the emergency department (ED) with upper respiratory symptoms, decreased appetite and 25 kg of unintentional weight loss. She was tachycardic with a heart rate of 119 bpm and temperature of 36.3°C; otherwise, her vital signs were within normal limits on room air, and she appeared fatigued and pale. She had primary respiratory acidosis with secondary metabolic alkalosis, leukocytosis with neutrophilia, microcytic anemia, and thrombocytosis, with WBC, hemoglobin, MCV, platelets, INR, and ferritin at 21, 114, 70, 476, 1.3, and 545 µg/L respectively. She also had normal TSH, creatinine, estimated GFR, urinalysis, rheumatoid factor, ANA and anti-ENA antibodies.

Her case is unique for multiple reasons. The pathology work-up was complicated, as mentioned above and detailed in the diagnostic assessment and discussion below. For cSCC, men are most affected, with patients over 60 years old comprising 80% of nonmelanoma skin cancers (1, 2). cSCC is uncommon in patients under 50 years old, and its recurrence or metastases usually occur within five years (not eight years) in 30–50% of patients (3). On the other hand, 85–90% of lung cancer cases can be attributed to smoking (4), with adults older than 65 most affected (5, 6) and men more affected than women (7). In non-small cell lung cancer (NSCLC), only 2–3% of patients have bilateral adrenal metastases at the initial presentation (8). Even if the biopsied adrenal mass represents an undifferentiated pleomorphic sarcoma (UPS), a type of high-grade aggressive soft tissue sarcoma (STS), only 10% of patients have detectable metastases at diagnosis (9). STS primarily affects adults older than 55 years (10). The patient's cardiac tumor extends from the right superior pulmonary vein (RSPV) into the left atrium (LA), which is rare and scarcely reported. One study examined 215 lung cancer patients with magnetic resonance angiography and found that two patients had a tumor that extended into the LA (11). A retrospective analysis of 4,668 patients who had surgery for lung cancer found evidence of pulmonary vein and LA involvement in 34 (0.7%) and 25 (0.5%) subjects (12, 13). Thus, our patient represents a minority of patients for any of these diseases. To our knowledge, this specific presentation has not been reported in the literature.

2 Case description

A non-smoker 38-year-old woman presented to the ED with upper respiratory symptoms, decreased appetite and 25 kg of

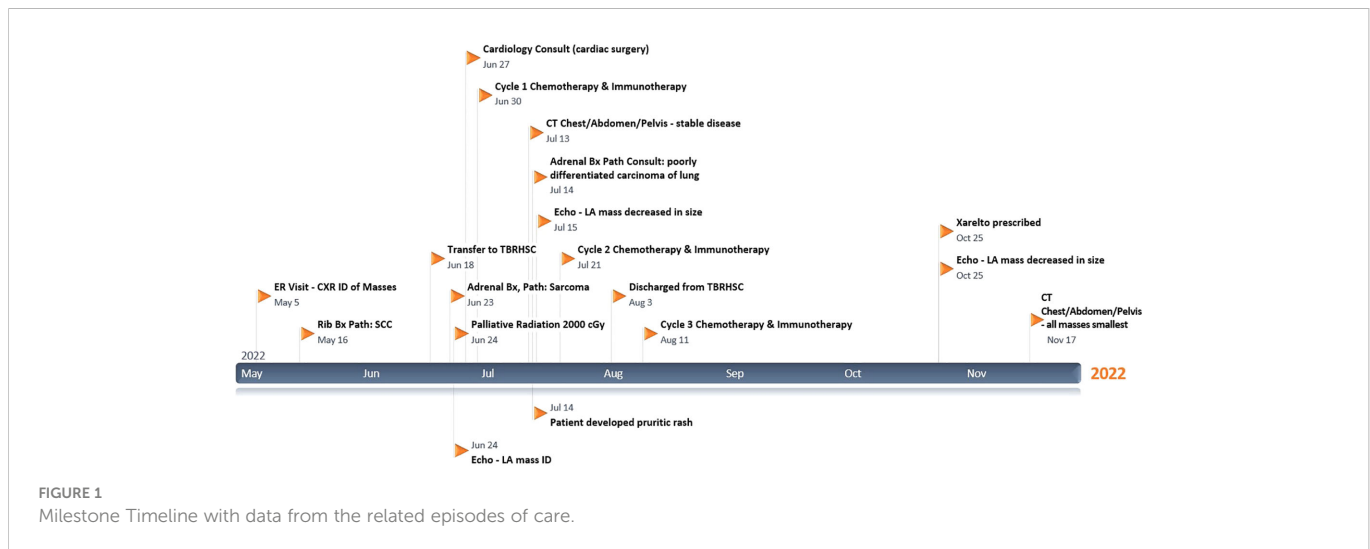
unintentional weight loss. Her past medical history was remarkable for cSCC of the posterior scalp, diagnosed in 2015. The tumor had been growing steadily for ten years. On March 2015, a head MRI showed a 7.5 x 4.8 x 7.0 cm posterior occipital midline extracalvarial mass. This moderately differentiated well-circumscribed posterior scalp cSCC was completely excised with no evidence of lymphovascular invasion. Besides previous intermittent asthma, she had no further history of medical conditions and did not use any medications. She had worked clerical jobs and, more recently, as a leather worker, which does not involve using chemicals. Her family history included a grandfather diagnosed with multiple myeloma. She was tachycardic with a heart rate of 119 bpm and temperature of 36.3°C; otherwise, her vital signs were within normal limits on room air, and she appeared fatigued and pale.

3 Clinical approach and timeline

Please refer to [Figure 1](#) for data from the related episodes of care.

3.1 Diagnostic assessment

The first investigation she had on presentation to the ED was, in fact, a chest X-ray. It showed a 5 x 4.5 cm right paratracheal mass, a 14 x 11 cm mass-like opacity obscuring the right cardiac border and right lower lung (RLL), and a 3.5 x 2.8 cm subtle left upper lung (LUL) opacity projected over the first rib. The following day, she had a thoracic CT with contrast which confirmed the 11.7 x 14.1 x 13.7 cm right middle lobe (RML) and RLL mass as highly suspicious for primary lung cancer. Additionally, it showed a nodular (1.7 cm) filling defect in the LA with direct extension through the RSPV and raised the concern of a filling defect in the right atrium extending through the inferior vena cava (IVC). Finally, it also showed mediastinal and hilar lymphadenopathy, including a 4.6 x 6.3 cm right paratracheal node slightly displacing the aorta and moderately compressing the superior vena cava (SVC), thyromegaly with a left-sided nodule and metastatic lesions to the left first rib (4.9 x 3.3 cm) and both adrenal glands. That same day, a neck and abdominopelvic CT confirmed the 3.7 x 3.2 x 5.2 cm left-lower thyroid nodule requiring FNA (but “unrelated to the presentation of lung mass”) and <1 cm right-sided thyroid nodules, as well as *osteitis condensans ilii*, splenomegaly, a 2 cm hepatic and a 2.9 cm ovarian cyst, a 3.5 cm posterior wall uterine



fibroid, and bilateral adrenal metastasis, measuring 8.2 x 6.0 x 3.2 cm (right) and 10.4 x 7.5 x 8.9 cm (left).

On May 13, she had ultrasound-guided biopsies of the left thyroid nodule (whose evaluation was limited by a low number of follicular cells) and the costochondral mass (which showed poorly differentiated SCC). Chest X-rays done on June 16 and June 18 showed that the multiple pulmonary masses were enlarging significantly, up to 16.4 cm (RLL), 6.8 cm (right paratracheal), and 6.2 cm (LUL). Head CT on June 19 showed no brain metastasis. An echocardiogram on June 24 estimated the baseline left ventricular ejection fraction (LVEF) at 70%, global longitudinal strain (GLS) -18%, and showed a 3.6 x 1.8 cm LA mass (from the RSPV) and an echogenic IVC (Eustachian) valve structure. Adrenal ultrasound on June 22 showed that the metastatic lesions in both glands were amenable for biopsy, hypovascular, and enlarging (12.3 x 9.1 x 6.9 cm on the right and 14.5 x 11.7 x 11.2 cm on the left). The biopsy of the left adrenal mass initially reported an UPS; however, a second pathologist reported it as a poorly differentiated carcinoma of lung origin, given the clinical history. Of note, at least three pathologists verified the specimens, including after being referred to a centre with an available pathology laboratory specializing in lung cancer. Diagnostic challenges included the adrenal biopsy being extensively necrotic and having little viable tissue available for examination and special stains. Still, immunohistochemistry was done in the biopsied specimens. The costochondral tumor was positive for P63 and negative for TTF-1 and PAX 8, excluding a primary adenocarcinoma and suggesting an SCC. It had a PD-L1 test with a 1-49% score, which guided management with Pembrolizumab as described below. The adrenal mass biopsy was negative for keratin, TTF-1, S100 protein, SOX10, smooth muscle actin (SMA), desmin, MDM2 and CDK4, and only vimentin was expressed. Occasionally SCCs can dedifferentiate and appear as spindle cell neoplasm. Multiple extrathoracic metastases make NSCLC stage IVB, and distant metastasis makes cSCC stage IV; however, studies in which most patients are like the one reported here and without previous systemic therapy are lacking to quantify her prognosis confidently.

3.2 Therapeutic intervention

On her first presentation, she was prescribed up to two doses of inhaled Salbutamol every 4 hours as needed. On the day of her first CT, she was started on Dexamethasone 4 mg PO to help reduce peritumoral edema. Before being transferred to Thunder Bay, she was started on Ceftriaxone 1 gr IV every 12 hours for a couple of days in case of an underlying infection. However, her leukocytosis was due to the dexamethasone use and extensive tumor burden. The Dexamethasone dose peak was 10 mg PO every 8 hours for three days during the first week of July, and by July 14, it was 4 mg PO twice a day up until discharge on August 2 with a plan to taper it over the next month. During this admission, she was also started on a Histamine-H₂-receptor antagonist (Famotidine 20 mg PO twice a day) while on steroids and Dextromethorphan 15 mg PO every 4 hours (antitussive) as needed.

In Thunder Bay, she had her urgent palliative radiotherapy to her mediastinum and right-hemithorax (2000 cGy in 5 fractions from June 20 until June 24), given her SVC and mainstem bronchus compression by the massive lung tumor. After it was completed, a multidisciplinary discussion involving oncology, cardiology and cardiac surgery determined that, based on the patient's performance status and disease extension, she would not tolerate cardiac surgery or a bronchoscopy, with very low chances of surviving without aggressive systemic therapy. The patient is aware of her limited second-line options if the chemo-immunotherapy fails.

Her systemic treatment was to be repeated every 21 days for four cycles. Each cycle included Pembrolizumab 2 mg/kg IV, a 50% dose-reduction of Paclitaxel (87.5 mg/m² IV) and kept the dose of Carboplatin down (380-400 mg IV) to minimize possible chemo-related side effects while providing the best possible treatment option to the patient. Due to personal and logistical challenges, her fourth cycle was 29 days instead of 21 days after the previous one. Three weeks after her fourth chemo-immunotherapy cycle, she started maintenance immunotherapy with Pembrolizumab 3.65 mg/kg IV on September 29. It will be repeated every 42 days until disease progression or unacceptable toxicity up to thirteen additional doses, whichever occurs first.

3.3 Follow-up and outcomes

Almost five months after her initial visit, her symptoms have significantly improved post-palliative radiotherapy and chemo-immunotherapy. Her Eastern Cooperative Oncology Group (ECOG) performance status has improved from grade 3 to 1. Follow-up diagnostic imaging includes post-radiotherapy thoracic CT in June and chest-abdominopelvic CT scans (Figure 2) in July and September 18, which show interesting results. Scans in June and July were consistent with continuously enlarging masses, as already mentioned, whose dimensions increased as follows: RML mass, 10 x 15 x 15–16.4 cm (including partial necrosis with vacuum phenomenon); right paratracheal node, 6.3 x 7.9 cm; left first rib, 6.4–7 x 5.7 cm; and bilateral adrenal metastasis (with some interval necrosis), measuring 10.7 x 9.8 x 9 cm (right) and 17.1 x 14.6 x 15.8 cm (left). Fortunately, CT on September 18 showed smaller lung (RML, 10 x 12.7 x 13.4 cm), nodal (right paratracheal, 5.5 x 5.6 cm), costochondral (left rib, 3.9 x 4 cm), and adrenal masses (right, 7.8 x 4.2 x 4 cm, and left, 18 x 12.6 x 12.2 cm), and increased necrosis. This CT scan also questioned a nonocclusive thrombosis in the IVC. A couple of repeat echocardiograms (Figure 3) in July (after chemo-immunotherapy cycles 1 and 2) confirmed the known LVEF, IVC echogenic structure and LA mass (2.9 x 2.3 cm).

Overall, she tolerated the therapeutic interventions above relatively well with an Edmonton Symptom Assessment System (ESAS) score <20, with maximum individual scores of 8 for tiredness and shortness of breath, which all improved concurrently with her ECOG performance status. On June 17,

before starting radiotherapy, she developed a paraneoplastic leukemoid reaction with WBC peaking at $67.3 \times 10^9/L$. The nadir of her anemia occurred one week after her first cycle of chemotherapy, with no overt hemorrhage or hemolysis and hemoglobin at 67 (for which she was transfused packed red blood cells). Her leukocytosis and anemia were resolved by September 26, with WBC and hemoglobin at 6.7 and 116, respectively. Unanticipatedly, she had developed a pruritic maculopapular rash on July 13 with unclear etiology, covering her face, especially lateral cheeks and forehead, neck, chest, upper arms, and groin. It was severe for at least a week, with thigh excoriations, arms plaques with minimal scaliness, and thoracic erythema; however, it subsided gradually with supportive care only. A dermatologist considered it a possible drug reaction likely secondary to Dextromethorphan more than Famotidine but did not think chemo-immunotherapy played a role in her eruptions.

Finally, her last echocardiogram (done on October 25) showed an LVEF >70%, a mildly reduced GLS at -18.7%, and the IVC content was again characterized as a thrombus. A recent CT scan in November and the abovementioned echocardiogram confirmed that the LA mass had reduced in size (2.5 x 1.9 cm). However, she has had palpitations, cough, and chest pain since this last month, feeling weaker. After discussing the echocardiogram findings with a Cardiologist, the Internal Medicine team prescribed her Rivaroxaban. Her November CT scan showed a large new right-sided pleural effusion and patchy consolidations deemed possibly due to post-radiation changes. Fortunately, it showed overall interval improvement. The stable IVC thrombus and the lung (RML, 9.5 x 11.3 x 12.5 cm), nodal (right paratracheal, 4.7 x 5 cm), and adrenal masses (right, 5.1 x 3 cm, and left, 10.3 x 9 cm) appeared the smallest they have been.

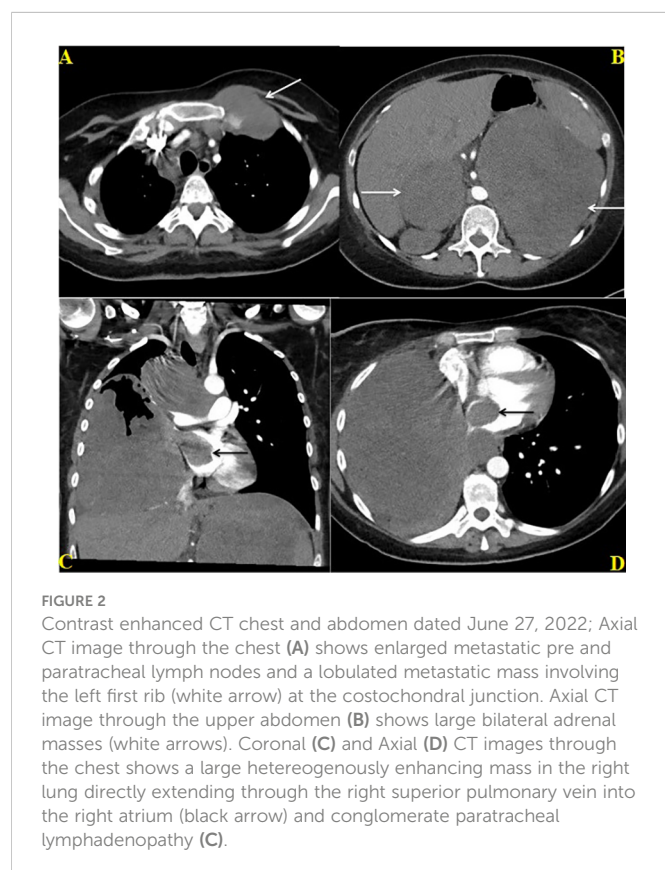


FIGURE 2
Contrast enhanced CT chest and abdomen dated June 27, 2022; Axial CT image through the chest (A) shows enlarged metastatic pre and paratracheal lymph nodes and a lobulated metastatic mass involving the left first rib (white arrow) at the costochondral junction. Axial CT image through the upper abdomen (B) shows large bilateral adrenal masses (white arrows). Coronal (C) and Axial (D) CT images through the chest shows a large heterogeneously enhancing mass in the right lung directly extending through the right superior pulmonary vein into the right atrium (black arrow) and conglomerate paratracheal lymphadenopathy (C).

4 Discussion

A strength of this report is that the case represents a frequent problem presenting uncommonly. A multidisciplinary approach, considering clinical presentation, medical imaging, and response to therapy (as assessed clinically and by imaging tests), can overcome diagnostic and therapeutic uncertainties that can occur with challenging pathology. Our approach's main limitations are those associated with pathology investigations. The patient presented like a textbook lung malignancy, with a sizeable pulmonary mass and metastases to the adrenals and bone. However, upon pathological inspection, this was more complex. Ultrasound-guided biopsy of the rib mass showed poorly differentiated SCC, and the left adrenal one was first considered a UPS, then a poorly differentiated carcinoma of lung origin. The uncertainty of this malignancy might be remedied with a lung biopsy. However, as Cancer Care Ontario suggests, if the patient has stage IV lung cancer, the least invasive and most accessible approach should be followed, which explains why only the rib and adrenal masses were biopsied (14).

One explanation for these differing diagnoses is that this patient has multiple primary malignancies. The first is the SCC which could be skin or lung, and a UPS (Figure 4). The lung tumor was not biopsied, so it is unclear whether there would be a third primary malignancy. The frequency of multiple primaries ranges from 2–17%.

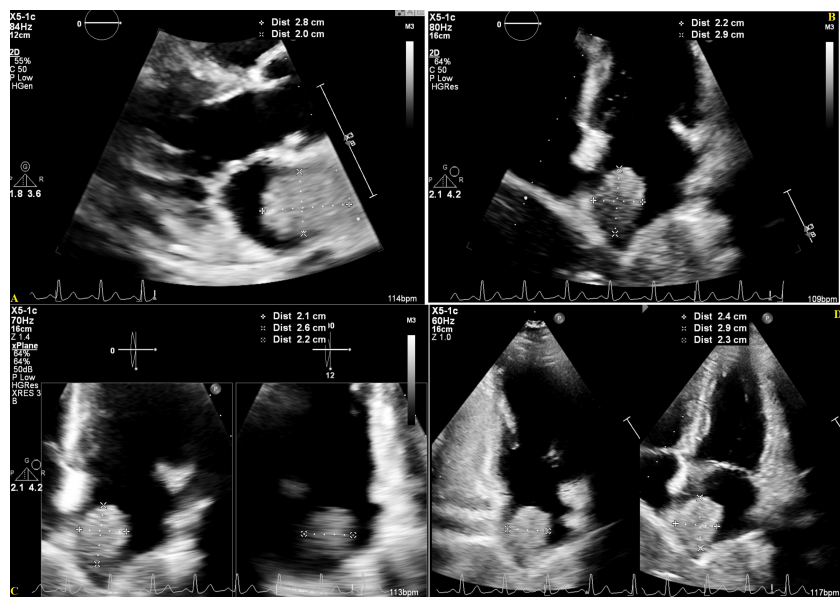


FIGURE 3
Transthoracic Echocardiogram showing a large echogenic mass in the Left Atrium, in the parasternal long-axis (A) and apical multi-chamber views (B–D), July 26, 2022, with measures estimating a mass size of 2.6–2.9 cm x 2.0–2.4 cm.

	Lung SCC	cSCC	Soft Tissue Sarcoma
Definition	Subtype of Non-Small Cell Lung Cancer, tumour cells have squamous appearance	Nonmelanoma skin cancer of keratinized epidermal keratinocytes	Very heterogeneous group of rare, solid cancers of mesenchymal or connective tissue origin
Etiology	Strong association with smoking, more so than other NSCLCs	Actinic damage due to chronic UV radiation exposure	Arise de novo with unclear etiology, Li-Fraumeni syndrome, neurofibromatosis type I. Additional factors: retinoblastoma, chemotherapy or radiation exposure, chronic irritation, lymphedema
Epidemiology	Leading cause of cancer death, men > 65 years of age more affected	Second most common nonmelanoma skin cancer, more common in men >60 years of age	<1% adult malignancies, median age at diagnosis: 65
Pathology	Solid, epithelial tumour; intercellular bridges; keratin pearls; expression of p40, p63	Keratinocytes, differentiation depends on grade	Desmin, vimentin, S100 antigen, neurofilaments
Clinical Presentation	Cough, dyspnea, chest pain, and hemoptysis; however, 25% patients may be asymptomatic, with an incidental finding on chest x-ray	Non-healing skin lesion usually found on the head or neck. Hard, scaly plaque with a poorly defined edge. A tumour may develop at the site of the lesion	Gradually enlarging, painless mass. Patients may describe paresthesia or edema in the extremity
Treatment	Surgical resection with/without chemotherapy	Surgical resection, radiation therapy, chemotherapy	Surgical resection, radiation therapy, chemotherapy
Prognosis	5-year survival rate: 20%. Better prognosis than SCLC	Excellent, with an overall five-year cure rate of >90 percent	≥ 33.3% will die from the disease, most due to lung metastases

FIGURE 4
Comparison of SCC of the lung, cutaneous SCC, and soft tissue sarcoma.

Cancer survivors may develop second primary malignancies due to an array of factors, including cancer predisposition syndromes or unique tumor characteristics, environmental exposures, and late effects of therapies (15, 16).

This patient may have a cancer predisposition syndrome, namely Li-Fraumeni syndrome. Li-Fraumeni syndrome is an inherited autosomal dominant disorder manifested by a wide range of malignancies that appear at an unusually early age (17). Li-Fraumeni syndrome is also known as Sarcoma, Breast, Leukemia, and Adrenal Gland cancer syndrome. This cancer predisposition syndrome is inherited as an autosomal dominant disorder and is associated with abnormalities in the tumor protein p53 gene (p53). Almost all types of soft tissue and bone sarcomas are seen in families with Li-Fraumeni syndrome. Patients with an abnormality in p53 who develop cancer are at a significantly increased risk of developing a second malignancy (18). For women, the lifetime risk of cancer approaches 100% and has been estimated to be ~90% by age 60 (19).

These differences in pathology can also be explained by tumor heterogeneity. Tumors can differ within a tumor (intra-tumoral heterogeneity) and between tumors (inter-tumoral heterogeneity). With intra-tumoral heterogeneity, histological examination often shows notable differences in the morphology of cancer cells in different areas of the same lesion, and areas of necrosis may be present. Also, the cells in a tumor may be cycling or non-cycling, quiescent, or reproductively dead (20). The stage of the cell cycle can influence cellular properties such as membrane biochemistry (21) and the ability to metastasize (22). Also, growth patterns, differentiation and keratinization grades between the primary and metastatic tumors can differ (23).

Lastly, it is also essential to consider the neglect of the first malignancy. The patient neglected the first cSCC for ten years and did not share with us why she waited so long to seek medical attention. Denial is a coping mechanism to adapt to stressful situations (24–26). It can become maladaptive when it leads to behaviors that are detrimental to their care, e.g., delaying seeking treatment. Many psychological, physiological, and demographic factors are present in those who show maladaptive denial (27), such as poor medical literacy, socioeconomic stresses, and tumors that grow slowly, as in our patient's case (28). Neglecting this high-risk malignancy might have had a role in cancer recurrence, categorized as such due to its size and invasion of the aponeurosis. Since no studies have been conducted with similar populations, the prognosis for the patient is unclear, especially if immunotherapy does not succeed. The patient would have limited options beyond a palliative approach in this case.

Clinical presentation and pathology can be quite different. Despite the pathology challenges, the patient responded well to the Paclitaxel-Carboplatin-Pembrolizumab regimen as the first-line treatment for metastatic SCC. The patient was treated accordingly based on clinical presentation and a multidisciplinary decision. Imaging studies confirmed the stabilization and improvement of her disease, already

hinted clinically with the resolution of her dyspnea and cough. This treatment resulted in stable disease and lessened symptom burden. Tumors may be heterogenous, and immunohistochemistry markers are not 100% sensitive. Thus, medical imaging is essential in directing metastatic cancer diagnosis, treatment, and follow-up, especially when pathology cannot confirm but only presume a specific diagnosis.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

All the listed authors contributed substantially to the conception or design of the work or the acquisition, analysis, or interpretation of data for the work. They drafted the work or revised it critically for important intellectual content and provided approval for publication of the content. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Firnhaber JM. Diagnosis and treatment of basal cell and squamous cell carcinoma. *Am Fam Physician* (2012) 86:161–8.
2. Diffey BL, Langtry JAA. Skin cancer incidence and the ageing population. *Br J Dermatol* (2005) 153:679–80. doi: 10.1111/J.1365-2133.2005.06799.X
3. Schmuls CD, Blitzblau R, Aasi SZ, Alam M, Andersen JS, Baumann BC, et al. *NCCN guidelines version 2.2022 squamous cell skin cancer* (2022). Available at: <https://www.nccn.org/home/member->.
4. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, et al. *NCCN guidelines version 5.2022 non-small cell lung cancer continue NCCN guidelines panel disclosures* (2022). Available at: <https://www.nccn.org/home/member->.
5. Planchard D, Popat S, Kerr K, Novello S, Smit EF, Faivre-Finn C, et al. Metastatic non-small cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2018) 29:iv192–237. doi: 10.1093/annonc/mdy275
6. Novello S, Barlesi F, Califano R, Cufer T, Ekman S, Levra MG, et al. Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2016) 27. doi: 10.1093/annonc/mdw326
7. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* (2022) 72:7–33. doi: 10.3322/caac.21708
8. Karanikiotis C, Tentes AA, Markakidis S, Vafiadis K. Large Bilateral adrenal metastases in non-small cell lung cancer. *World J Surg Oncol* (2004) 2:1–7. doi: 10.1186/1477-7819-2-37/TABLES/1
9. WHO Classification of Tumours Editorial Board. *Soft tissue and bone tumours. 5th ed.* (Lyon, France: International Agency for Research on Cancer Press) (2020).
10. Nystrom LM, Reimer NB, Reith JD, Dang L, Zlotecki RA, Scarborough MT, et al. Multidisciplinary management of soft tissue sarcoma. *Sci World J* (2013) 2013. doi: 10.1155/2013/852462
11. Takahashi K, Furuse M, Hanaoka H, Yamada T, Mineta M, Ono H, et al. Pulmonary vein and left atrial invasion by lung cancer: Assessment by breath-hold gadolinium-enhanced three-dimensional MR angiography. *J Comput Assist Tomogr* (2000) 24:557–61. doi: 10.1097/00004728-200007000-00008
12. Stella F, Dell'Amore A, Caroli G, Dolci G, Cassanelli N, Luciano G, et al. Surgical results and long-term follow-up of T4-non-small cell lung cancer invading the left atrium or the intrapericardial base of the pulmonary veins. *Interact Cardiovasc Thorac Surg* (2012) 14:415–9. doi: 10.1093/icvts/ivr160
13. Riquet M, Grand B, Arame A, Pricopi CF, Foucault C, Dujon A, et al. Lung cancer invading the pericardium: Quantum of lymph nodes. *Ann Thorac Surg* (2010) 90:1773–7. doi: 10.1016/J.ATHORACSUR.2010.07.039
14. Cancer Care Ontario. *Lung cancer imaging guidelines: Integration with the lung cancer diagnosis and staging clinical pathway* (2014). Available at: <https://www.cancercareontario.ca/en/guidelines-advice/types-of-cancer/3201>.
15. Coyte A, Morrison DS, McLoone P. Second primary cancer risk - the impact of applying different definitions of multiple primaries: Results from a retrospective population-based cancer registry study. *BMC Cancer* (2014) 14:1–11. doi: 10.1186/1471-2407-14-272
16. Vogt A, Schmid S, Heinemann K, Frick H, Herrmann C, Cerny T, et al. Multiple primary tumours: Challenges and approaches, a review. *ESMO Open* (2017) 2:e000172. doi: 10.1136/esmoopen-2017-000172
17. Malkin D. Li-Fraumeni syndrome. *Genes Cancer* (2011) 2:475–84. doi: 10.1177/1947601911413466
18. Bougeard G, Renaux-Petel M, Flaman JM, Charbonnier C, Fermey P, Belotti M, et al. Revisiting Li-fraumeni syndrome from TP53 mutation carriers. *J Clin Oncol* (2015) 33:2345–52. doi: 10.1200/JCO.2014.59.5728
19. Mai PL, Best AF, Peters JA, DeCastro RM, Khincha PP, Loud JT, et al. Risks of first and subsequent cancers among TP53 mutation carriers in the national cancer institute Li-fraumeni syndrome cohort. *Cancer* (2016) 122:3673–81. doi: 10.1002/cncr.30248
20. Heppner GH, Miller BE. Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer Metastasis Rev* (1983) 2:5–23. doi: 10.1007/BF00046903
21. Pasternak CA, Warmley AMH, Thomas DB. STRUCTURAL ALTERATIONS IN THE SURFACE MEMBRANE DURING THE CELL CYCLE. *J Cell Biol* (1971) 50:562. doi: 10.1083/jcb.50.2.562
22. Suzuki N, Withers HR, Koehler MW. Heterogeneity and variability of artificial lung colony-forming ability among clones from mouse fibrosarcoma. *Cancer Res* (1978) 38:3349–51.
23. Ichinose J, Shinozaki-Ushiku A, Takai D, Fukayama M, Nakajima J. Differential diagnosis between primary lung squamous cell carcinoma and pulmonary metastasis of head and neck squamous cell carcinoma. *Expert Rev Anticancer Ther* (2016) 16:403–10. doi: 10.1586/14737140.2016.1147352
24. Siemerink EJM, Jaspers JPC, Plukker JTM, Mulder NH, Hospers GAP. Retrospective denial as a coping method. *J Clin Psychol Med Settings* (2011) 18:65–9. doi: 10.1007/S10880-011-9223-X
25. Hackett TP, Cassem NH. Development of a quantitative rating scale to assess denial. *J Psychosom Res* (1974) 18:93–100. doi: 10.1016/0022-3999(74)90072-5
26. Vos MS, de Haes JCJM. Denial in cancer patients, an explorative review. *Psychooncology* (2007) 16:12–25. doi: 10.1002/PON.1051
27. Alam M, Goldberg LH, Silapunt S, Gardner ES, Strom SS, Rademaker AW, et al. Delayed treatment and continued growth of nonmelanoma skin cancer. *J Am Acad Dermatol* (2011) 64:839–48. doi: 10.1016/j.jaad.2010.06.028
28. Varga E, Korom I, Raskó Z, Kis E, Varga J, Oláh J, et al. Neglected basal cell carcinomas in the 21st century. *J Skin Cancer* (2011) 2011. doi: 10.1155/2011/392151



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

Salvatore Annunziata,
Head of TracerGLab (IRCCS), Italy
Ryogo Minamimoto,
National Center for Global Health and
Medicine, Japan

*CORRESPONDENCE

Xiaoliang Chen
✉ chenxiaoliang26@163.com
Yao Liu
✉ 648283926@qq.com

[†]These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 06 December 2022

ACCEPTED 10 January 2023

PUBLISHED 26 January 2023

CITATION

Liu Q, Yang T, Chen X and Liu Y (2023)
Clinical value of ¹⁸F-FDG PET/
CT in the management of
HIV-associated lymphoma.
Front. Oncol. 13:1117064.
doi: 10.3389/fonc.2023.1117064

COPYRIGHT

© 2023 Liu, Yang, Chen and Liu. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Clinical value of ¹⁸F-FDG PET/CT in the management of HIV- associated lymphoma

Qi Liu^{1†}, Tao Yang^{2†}, Xiaoliang Chen^{1*} and Yao Liu^{2*}

¹Department of Nuclear Medicine, Chongqing University Cancer Hospital, Chongqing, China,

²Department of Hematology-Oncology, Chongqing University Cancer Hospital, Chongqing, China

HIV is still a major public health problem. At present, HIV-associated lymphoma remains the leading cause of deaths among people living with HIV, which should be paid more attention to. ¹⁸F-fluorodeoxyglucose (FDG) PET/CT has been recommended in the initial staging, restaging, response assessment and prognostic prediction of lymphomas in general population. HIV-associated lymphoma is, however, a different entity from lymphoma in HIV-negative with a poorer prognosis. The ability to accurately risk-stratify HIV-infected patients with lymphoma will help guide treatment strategy and improve the prognosis. In the review, the current clinical applications of ¹⁸F-FDG PET/CT in HIV-associated lymphoma will be discussed, such as diagnosis, initial staging, response evaluation, prognostic prediction, PET-guided radiotherapy decision, and surveillance for recurrence. Moreover, future perspectives will also be presented.

KEYWORDS

HIV, diffuse large B-cell lymphoma, Burkitt's lymphoma, Hodgkin's lymphoma, ¹⁸F-FDG PET

1 Introduction

Acquired immunodeficiency syndrome (AIDS) is a major disease that greatly endangers human health across the world. AIDS is caused by human immunodeficiency virus (HIV), which progressively destroys CD4 + T cells and compromises the function of the immune system leading to severe opportunistic infection or cancer (1). HIV infection is associated with an increased risk of a range of cancers, which are classified into two groups: AIDS-defining cancers (ADCs) and non-AIDS-defining cancers (NADCs) (2). There are three ADCs: Kaposi sarcoma (KS), non-Hodgkin lymphoma (NHL), and cervical cancer. The incidence of NHL exceeded KS in 2018 and became the highest ADC according to the statistics of American National Cancer Institute, which was 1194 and 765 cases/year, respectively (2). With the exception of the three ADCs, all cancers in HIV-infected population are considered as NADCs, including HIV-associated tumors and incidental cancers. The most common types of NADCs found in persons living with HIV include those with a higher rate among HIV-infected population (such as anal cancer, liver cancer, oropharyngeal cancer, Hodgkin lymphoma and lung cancer, which compose about half of all cancers among HIV-infected persons in the United States) and those common in the general population (such as prostate, breast, and colon

cancer) (2, 3). In a meta-analysis of 4,797 NADC cases, the most common NADC type was lung cancer (17.6%), followed by Hodgkin lymphoma (13.4%), anal cancer (5.3%) and liver cancer (3.5%) (4). With the advent of the highly active antiretroviral therapy (HAART) era, the standard incidence rate of NHL has dropped dramatically, whereas the standard incidence rate of HL increased (5, 6). At present, HIV-associated lymphoma remains the leading cause of deaths among people living with HIV, which should be paid more attention to (7, 8).

Diffuse large B-cell lymphoma (DLBCL) and Burkitt's lymphoma (BL) are among the most common subtypes of NHL in ADC in the HAART era (2, 9), with an incidence of 1.20 and 0.32 per 1000 person-years, respectively (9, 10). Other rare forms of NHL, including plasmablastic lymphoma (PBL), primary effusion lymphoma (PEL), and primary central nervous system lymphoma (PCNSL), are also associated with HIV infection (11–13). Furthermore, compared to HIV-negative persons, the risk of HL which is not AIDS-defining is higher among HIV-infected persons (3).

Previous studies have shown that ^{18}F -fluorodeoxyglucose (FDG)-positron emission tomography (PET)/computed tomography (CT) plays an important role in management of HIV-negative patients with lymphoma (14, 15). This noninvasive technology provides several quantitative metabolic parameters such as the maximum of standard uptake value (SUV_{max}), SUV at lean body mass (SUL), metabolic tumor volume (MTV), and total lesion glycolysis (TLG) for the management of lymphoma. Due to high sensitivity to whole-body involved lymph nodes and extranodal lymphoma, ^{18}F -FDG PET/CT has been incorporated into standard assessment for all FDG-avid lymphomas, and is recommended by the 2014 Lugano criteria (16). Meanwhile, several large randomized clinical trials demonstrated that, decisions to reduce or increase the intensity of treatment can be based on the changes in tumor metabolism during lymphoma treatment (17–19).

However, HIV-associated lymphoma is a different entity from lymphoma in the general population. Compared with lymphoma in HIV-negative persons, HIV-associated lymphoma has a worse prognosis, more frequent occurrence of B symptoms (i.e., weight loss, night sweats, and fever) and more advanced Ann Arbor stages, even in the HAART era (20, 21). HIV-related benign lymphadenopathy is a main reason for a false positive of HIV-associated lymphoma on FDG PET. Whether HIV-infected status will increase tumor burden and reduce overall survival is still under discussion. To date, there are relatively few studies on the clinical value of FDG PET in HIV-associated lymphoma.

This article presents an overview of current ^{18}F -FDG PET/CT clinical applications in the management of patients with HIV-associated lymphoma, including initial staging, response evaluation, prognostic prediction, PET-guided radiotherapy decision and differentiated diagnosis. Furthermore, future perspectives will also be discussed.

2 PET to distinguish lymphoma from benign lesions

2.1 HIV-associated lymphoma vs. benign lymphadenopathy

Lymphadenopathy in HIV-infected patients is a diagnostic challenge for nuclear medicine physicians and clinicians. HIV has a

strong tropism for lymphoid tissues, and it leads to a progressive alteration in lymph nodes architecture that clinically manifests as persistent lymphadenopathy in three histologic phases: stage I, explosive follicular hyperplasia with large germinal centers; stage II, germinal center involution and lymphocytes depletion; stage III, proliferation of blood vessels inside the lymph nodes (22). In addition to the HIV infection itself, the other main causes of lymphadenopathy include HIV-associated lymphoma, opportunistic infections, and inflammatory conditions (23–25). Immune reconstitution inflammatory syndrome (IRIS) should also be considered as a possible cause of lymphadenopathy in patients who start HAART. The incidence of lymphadenopathy is negatively correlated with CD4 counts (26). Due to the different respective treatments and adverse effects of over treatment, it is vital to distinguish HIV-associated lymphoma from benign lymphadenopathy.

A qualitative analysis assessed the role of PET/CT in differentiating NHL from benign lymphadenopathy in patients with HIV-1. In 16 PET/CT scans, PET/CT accurately depicted the context of lymphoma. 12 patients with concordant PET/CT (+/+) findings had well-suppressed viral loads and low CD4 levels, whereas the 4 PET/CT (+/–) scans had the highest values of both laboratory parameters. In the case of inconsistent PET and CT findings, increased viral loads and CD4 count may imply benign lymphadenopathy (27). Since the advent of HAART, PET/CT (+/–) findings concurrent with a rapid increase of CD4 count may imply benign lymphadenopathy, perhaps due to IRIS (28, 29). In acute phase of HIV, FDG-avid lymph nodes are typically observed in the head and neck. In late phase, involvement of abdominal lymph nodes is more common (30). In addition, Liu et al. showed that concurrent nasopharyngeal lesions and lymphadenopathy on PET/CT had a greater possibility of malignant lymphoma instead of benign lymphoproliferative disease or nasopharyngeal carcinoma (31).

In quantitative analysis of ^{18}F -FDG PET/CT, Chen et al. identified that the SUV_{max} of only lymph nodes (SUV_{LN}), and the most FDG-avid lesion to liver SUV_{max} ratio (SUR_{max}) provided a new basis for distinguishing malignant lymphoma from inflammatory lymphadenopathy in patients with HIV. This study retrospectively assessed 59 HIV-infected patients, of which 37 had HIV-associated lymphoma, and 22 had HIV-associated inflammatory lymphadenopathy. Malignant lymphoma invaded more commonly to extra-lymphatic lesions, compared to inflammatory lymphadenopathy (83.8% vs. 54.5%, $p=0.000$). Especially, the involved lesions of digestive tract and Waldeyer's ring obviously differed between the two groups ($p=0.004$ and 0.033 , respectively). Furthermore, the SUR_{max}, SUV_{LN}, SUV_{Marrow} and SUV_{Liver} in malignant lymphoma were significantly higher than those in inflammatory lymphadenopathy ($p=0.000$, 0.000 , 0.002 and 0.017 , respectively). The cut-off point of 3.1 for SUR_{max} showed the best equilibrium between sensitivity (68.2%) and specificity (91.9%), and the cut-off point of 8.0 for the SUV_{LN} had also high specificity (89.2%) and relatively reasonable sensitivity (63.6%). The cut-off point of 5 for numbers of involved areas and 3.6 centimeter for maximum diameters of lymph node had relatively low sensitivity (62.2%, 64.9%, respectively) and specificity (72.7%, 86.4%, respectively) (32).

Another retrospective study assessed the diagnostic accuracy of the quantitative PET indices for distinguishing HIV-associated lymphoma from reactive lymphadenopathy, including MTV, TLG, maximum and peak of SUL (SUL_{max} and SUL_{peak}). All of the

quantitative PET metrics were significantly higher in patients with lymphoma than in those with reactive lymphadenopathy (all *p* values, <0.001). The cut-off points of 173 for TLG and 53.8 for MTV showed high sensitivity (89%, 84%, respectively) and specificity (100%, 100%, respectively) for differentiating lymphoma from reactive lymphadenopathy. A summed SULpeak cut-off point of 23.8 showed a sensitivity and specificity of 84% and 95% respectively, whereas a summed SULmax cut-off point of 28.4 yielded a sensitivity and specificity of 84% and 82% respectively. In the reactive lymphadenopathy group, PET metrics were positively correlated with viral load. However, in the HIV-associated lymphoma group, none of the PET parameters showed a significant correlation with viral load. In the qualitative MIP symmetry score analysis of the pattern of lymph node involvement, asymmetrical FDG uptake had an accuracy of 90.4% for differentiating lymphoma from reactive adenopathy in HIV-infected patients (33). The diagnostic efficacies of FDG PET/CT PET/CT in differentiating HIV-associated lymphoma from benign lymphadenopathy are summarized in Table 1.

2.2 PCNSL vs. cerebral opportunistic infection

PCNSL and cerebral toxoplasmosis (CTOX) are the two most common diagnoses of intracranial mass in HIV-infected patients. PCNSL is a rare extranodal NHL, with an incidence of 4-5/1000 among patients with HIV (34). The pooled prevalence of CTOX among HIV-infected patients is approximately 35.8% (35). It is still a challenge for conventional imaging to distinguish these two entities because of overlapping and non-specific imaging features (35, 36).

Because of the poor prognosis and the fatal outcome of misdiagnosis or delayed treatment (37), accurate and prompt diagnosis is crucial.

In a prospective study of HIV-infected patients, Westwood et al. identified that foci of increased FDG uptake corresponding to the enhanced lesions on MRI was considered suggestive of PCNSL. Conversely, no clinically significant increase in FDG uptake was considered suggestive of CTOX. One patient had progressive multifocal leukoencephalopathy (PML) with equivocal metabolic activity, and the other one with hemorrhagic brain metastasis showed normal metabolic activity (38).

Hoffman et al. showed increased FDG uptake of PCNSL compared with benign lesions. For qualitative analysis, a visual scoring system was proposed. For semiquantitative assessment, a count ratio of lesion to contralateral homologous brain was calculated. Both qualitative visual scoring and count ratios (1.8, 0.65, 1.3, and 0.7, respectively; *p*=0.006) suggested higher scores and count ratios in lymphoma than those in nonmalignant lesions (toxoplasmosis, PML, syphilis) (39). These results were supported by another semiquantitative study, in which the SUV ratio of the lesion to the contralateral brain in HIV-infected patients with PCNSL was higher than that in patients with cerebral infections including toxoplasmosis and tuberculoma (range, 1.7-3.1, 0.3-0.7, respectively; *p*<0.05) (40). A similar semiquantitative analysis suggested that PCNSL showed increased lesional uptake in comparison with the normal brain cortex (mean SUVmax, 18.8; range, 12.4-29.9), while CTOX had lesional uptake less than that of normal brain cortex (mean SUVmax, 3.5; range, 1.9-5.8) (41).

It should be noted that the challenge in applying semiquantitative analyses similar to the above methods arises when the corresponding contralateral brain exhibits physiologically increased FDG uptake.

TABLE 1 The diagnostic efficacies of FDG PET/CT in differentiating HIV-associated lymphoma from benign lymphadenopathy.

References	Number of patients	Subtypes of lymphoma	Parameters	Cut-off point	AUC	P value	Specificity (%)	Sensitivity (%)	PPV (%)	NPV (%)
32	59 (37 HAL)	35 B-cell lymphoma, 1 HL, 1 T-cell lymphoma	SURmax	3.1	0.888	0.000	91.9	68.2	NA	NA
			SUV _{LN}	8.0	0.815	0.000	89.2	63.6	NA	NA
			SUV _{Marrow}	NA	0.611	0.156	NA	NA	NA	NA
			SUV _{Liver}	NA	0.567	0.393	NA	NA	NA	NA
			Number of lymph node involved areas	5	0.692	0.014	72.7	62.2	NA	NA
			Maximum diameter of lymph nodes	3.6	0.768	0.001	86.4	64.9	NA	NA
33	41 (19 HAL)	16 DLBCL, 3 HL	Single SULmax	7.8	0.971	NA	100	89	100	92
			TLG	173.0	0.964	NA	100	89	100	92
			Single SULpeak	6.6	0.964	NA	100	84	100	88
			MTV	53.8	0.957	NA	100	84	100	88
			Sum SULpeak	23.8	0.935	NA	95	84	94	88
			Sum SULmax	28.4	0.904	NA	82	84	80	86

Statistically significant values are highlighted in bold.

1. AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; NA, not available; 2. DLBCL, diffuse large B-cell lymphoma; HL, Hodgkin's lymphoma; 3. SUVmax, the maximum of standard uptake value; SURmax, liver SUVmax ratio; SUV_{LN}, SUVmax of only lymph nodes; SUL, SUV at lean body mass; SULmax, the maximum of SUL; SULpeak, the peak of SUL; TLG, total lesion glycolysis; MTV, metabolic tumor volume.

Nevertheless, delayed ^{18}F -FDG PET/CT has been found to be effective in differentiating malignancy from inflammation or infection. Compared to benign lesions or normal tissues, malignant tumors show a progressive increase in metabolic activity, which have improved the background contrast with lesion-based sensitivity as high as 98% (42, 43).

3 PET for initial staging

As described above, HIV-positive patients are at higher risk of lymphoma, making the accurate appreciation of the extent of disease based on reliable initial staging imperative. The role of ^{18}F -FDG PET/CT in initial staging of lymphoma is superior to contrast-enhanced CT, especially for the detection of lesions with no or minor anatomical abnormalities (including bone marrow, spleen and gastrointestinal tract involvement) (44, 45). Due to that PET scans may detect additional disease sites, the clinical stage is modified in 15% to 20% of patients and therapeutic decision is changed in 8% of patients (46).

Just et al. retrospectively studied 13 HIV-infected patients with BL who underwent one or more PET/CT scans (47). In 5 of 5 patients scanned before treatment, PET/CT demonstrated all involved sites detected at conventional imaging and identified additional sites with high SUVmax in 4 of 5 patients, lymph nodes mainly (4/5), and also spleen (1/5), bone (1/5), and peritoneum (2/5). Lymph node involvement was found in 54% of patients, which is known to be unusual in endemic or sporadic BL. Besides, in 3 patients, BL was predominantly located in the parotid lymph nodes, which is also not a common finding.

As mentioned above, it is a challenge to distinguish HIV-associated lymphoma from benign lymphadenopathy. The likelihood of false positives due to benign lymphadenopathy should be fully taken into consideration when clinicians use FDG PET/CT for initial staging of HIV-associated lymphoma. Patients' clinical presentations, CD4 counts, imaging features and PET metabolic parameters are helpful to identify the possible cause of lymphadenopathy. However, there are very limited studies on ^{18}F -FDG PET/CT for initial staging of HIV-associated lymphoma to date, with small sample sizes. Prospective studies with larger samples are needed in the future.

4 PET for response evaluation and prognostic prediction

^{18}F -FDG PET/CT has been identified as a reliable technique for both initial staging and early response evaluation in lymphoma, making individualized patient management possible. Previous multicenter studies have found that stage- and risk-adapted treatment is effective and feasible (48, 49). Furthermore, as the 5-year survival rates of both HL and NHL improving, the current interest has shifted towards reducing treatment-related complications (e.g. secondary malignant tumor and cardiovascular events) (50). The ability to promptly and accurately risk-stratify HIV-infected patients with lymphoma will improve the prognosis and decrease the adverse effects of overtreatment.

4.1 Baseline PET/CT

The metabolic parameters including SUV, MTV and TLG on baseline PET have been considered as useful predictors of tumor aggressiveness and response to treatment in solid tumors (16, 51). However, the current literatures of HIV-associated lymphoma are somewhat conflicting. Louarn et al. found that all the parameters on baseline PET were associated with progression-free survival (PFS) in univariate analysis, whereas a high total MTV was the only metabolic parameter independently correlated with PFS (Hazard ratio, 3.62) in multivariate analysis. The optimal total MTV cut-off point for prognostic assessment was 527 cm^3 , with a 2-year PFS of 71% ($>527\text{ cm}^3$) compared with 91% ($\leq 527\text{ cm}^3$) with p value of 0.004 (52).

On the contrary, Lawal et al. found that the SUVmax, MTV and TLG of lesions were not significantly different between the HIV-positive and HIV-negative groups, whereas presence of HIV infection was related with significantly higher rate of treatment failure following the ABVD (Adriamycin, bleomycin, vinblastine and dacarbazine) regimen compared with the non-infected group (40.4%, 17.7%, respectively, $p=0.0034$). In univariate analysis, only HIV status of patients was a significant predictor of therapy outcome ($p<0.001$). SUVmax, MTV and TLG of lesions as well as the Ann Arbor stage were not significant in predicting therapy outcome. Furthermore, a multiple logistic regression showed that HIV status alone was significant in predicting therapy outcome [overall rate (OR) = 2.930, 95% confidence interval (CI), 1.197-7.172, $p=0.023$] among the metabolic parameters and Ann Arbor stage (53).

These findings were supported by another study by the same author. Lawal et al. analyzed 160 patients with HL, including 57 HIV-positive patients. The median values of SUVmax, SUVmean, MTV and TLG were slightly higher among the HIV-positive group compared with the HIV-negative group. However, no significant difference was found between the two groups. In addition, among the seven parameters of International Prognostic Score (IPS) indicating poor prognostic factors, only male sex (HIV-negative group higher, $p=0.005$) and serum albumin less than 4 g/dl ($p=0.009$) were found with significant differences between the two groups (54).

In paediatric and adolescent patients with HIV-associated lymphoma, similar results were found in a retrospective study. Reed et al. found that HIV status and treatment response on PET were significantly related to PFS ($p=0.036$, $p<0.001$, respectively) in univariate analysis. In contrast, none of the metabolic parameters was significantly predictive of either PFS or overall survival (OS). Unlike in the adult research, the baseline total MTV was considered as a significant predictor of treatment response ($p=0.017$) (55).

4.2 Interim PET/CT

Interim ^{18}F -FDG PET/CT in both HL and NHL has demonstrated prognostic significance in the general population: residual tumor avidity on PET after 2 to 4 cycles corresponds with lower PFS, while interim PET negative is correlated with higher PFS (56–58). Early identification of patients at high risk of standard

treatment failure will affect treatment decisions (more intensive therapy) and ultimately improve patient treatment outcomes and survival.

However, sparse studies exist about the prognostic value of interim ^{18}F -FDG PET/CT in patients with HIV-associated lymphoma. Okosun et al. showed that interim PET negativity was significantly associated with a higher PFS. A total of 23 patients with advanced HIV-associated HL were included in this analysis. Deauville criteria was used to evaluate the metabolic activity of lesions, and a Deauville score of 1-3 was defined as PET-negative. After a median follow-up of 27 months (range, 12-50 months), the 2-year PFS rates for interim PET-positive and PET-negative patients were 50% and 100% respectively (log-rank $p=0.0012$) (59).

A similar result of relationship between interim PET findings and OS was reported by Minamimoto, R. et al. (60). A total of 24 patients with HIV-related malignant lymphoma (13 DLBCL, 11 BL) were included in this study, who underwent interim PET/CT. In 10 of 24 cases, interim PET findings were evaluated as “positive”, while the rest cases were evaluated as “negative”. Interim PET negativity was associated with significantly longer OS (932 ± 549 days) compared to positive cases (454 ± 442 days, $p=0.043$). Over all two year survival rate of negative findings on interim PET was 80% (95%CI, 69%-91%), which was higher than 29% (95%CI, 16%-41%) in positive cases. Moreover, Cox regression analysis showed strong prognostic influences of interim PET findings (Hazard ratio 4.57, 95%CI 0.88-23.73, $p=0.07$) and Eastern Cooperative Oncology Group performance status (Hazard ratio 10.52, 95%CI 1.26-87.82, $p=0.03$) on OS.

A phase 2 trial implied that response-adapted therapy based on interim PET was feasible in patients with HIV-associated HL (61). All patients underwent interim PET scan after 2 initial cycles of ABVD, and 10 of 12 patients achieved PET-negative status, while 2 of 12 remained PET-positive (Deauville scores 4 or 5) according to Deauville criteria. All the PET-negative and one PET-positive patients continued ABVD regimen. The other PET-positive patient received 6 cycles of BEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone) regimen. Finally, 75% of HIV-HL patients achieved complete response (CR) and 25% partial response (PR). With a median follow-up of 39 months (range, 5-53 months), a 2-year PFS was 83% (95% CI, 46.1%-95.3%), which was similar to that in non-HIV-HL patients in the same phase 2 trial (2-year PFS, 79%) (62). The prognostic predictions of both baseline and interim FDG PET for HIV-associated lymphoma are summarized in Table 2.

5 PET-guided radiotherapy decision

The success of an individualized treatment decision depends largely on accurate staging and response assessment. Engert A et al. have suggested that PET done after chemotherapy was helpful to guide the need for additional radiotherapy in patients with HL lymphoma (63). In advanced-stage HL, patients with a persistent mass after chemotherapy measuring 2.5 cm or larger and positive on PET scan should receive additional radiotherapy. For patients with

nonbulky stage I/II DLBCL and negative on interim PET scan, chemotherapy alone was comparable to chemotherapy followed by radiotherapy; additional radiotherapy was applied to patients with positive on interim PET scan (64, 65).

6 PET at the end of therapy

FDG-PET plays an important role in the response assessment of both HL and NHL at the end of therapy, especially in the identification of a CT-detected residual mass (16). An FDG-PET/CT scan after completion of the intended treatment can differentiate viable tumor cells from fibrosis or necrosis. Engert A et al. reported that post-treatment PET scans were able to sharply reduce the number of patients with additional radiotherapy for residual mass to 11% from 71% in previous trials (66). The Deauville 5-point scale is used to assess differing degrees of response at the end of treatment (67). A Deauville score of 1 to 3 identifies complete metabolic response, while residual disease is defined by a score of 4 to 5.

7 Surveillance PET scans

Relapses occur in approximately 10-20% of patients with early stage HL, as well as 30-40% in those at advanced stages following first-line therapies (68). There is sparse available published literature of relapsed HIV-associated lymphoma. In general population, the role of FDG PET in post-treatment surveillance remains controversial.

In a prospective study, 5 patients were detected relapsed or refractory disease by FDG PET, while 6 patients were found to have false-positive findings (69). These findings are corroborated by a number of other studies. A multicenter retrospective study of 161 patients with HL showed that the overall positive predictive value (PPV) and negative predictive value (NPV) of PET were 28% and 100% respectively (70). In a retrospective analysis of 75 patients with DLBCL, the PPV of PET was 85%, but usefulness was only for high-risk patients with symptoms indicative of a relapse and those older than 60 years (71).

Thus, there is currently insufficient evidence to recommend PET as a routine surveillance tool for patients with lymphoma. FDG PET should be used only for patients at a high-risk of recurrence to reduce radiation burden and costs.

8 A simple comparison of the role of FDG PET in COVID-19 and HIV pandemic

A systemic review was conducted to evaluate the role of ^{18}F -FDG PET/CT in patients with Coronavirus Disease (COVID-19) (72). In 10 of 11 studies, FDG PET/CT was used to assess the oncological indications and incidental findings suspicious of COVID-19 infection, particularly in suspected COVID-19 interstitial pneumonia. Only in one study, FDG PET/CT was performed to

TABLE 2 The prognostic predictions of both baseline and interim FDG PET for HIV-associated lymphoma.

PET imaging	References	Study type	Number of patients	Subtypes	Parameters	Cut-off value	P value	AUC	Complete remission (%)	Survival (%)
Baseline PET	52	Retrospective	109, HIV, +	HL	tMTV	≤527	0.004	0.60	87	2-year PFS 915-year OS 89.1
						>527				2-year PFS 715-year OS 69.4
					TLG	≤230	NS	0.59		5-year OS 96.0
						>230				5-year OS 82.9
					SUVmax	≤8.7	NS	0.55		5-year OS 94.4
						>8.7				5-year OS 81.1
					SUVpeak	≤7.1	NS	0.55		5-year OS 94.6
						>7.1				5-year OS 80.8
					SUVmean	≤5.1	NS	0.53		5-year OS 94.3
						>5.1				5-year OS 81.1
	53	Prospective	136 (57, HIV, +)	HL	SUVmax	NA	0.087	NA	NA	NA
					MTV	NA	0.589	NA		NA
					TLG	NA	0.460	NA		NA
	54	Retrospective	160 (57, HIV, +)	HL	SUVmax	NA	0.401	NA	NA	NA
					SUVmean	NA	0.056	NA		NA
					MTV	NA	0.766	NA		NA
					TLG	NA	0.965	NA		NA
	55	Retrospective	69 (13, HIV, +)	HL	SUVmax	NA	0.379	NA	58	NA
					MTV	NA	0.065	NA		NA
					TLG	NA	0.099	NA		NA
Interim PET	59	Retrospective	23, HIV, +	HL	PET negative	NA	0.0012	NA	96	2-year PFS 100
					PET positive	NA				2-year PFS 50
	60	Retrospective	24, HIV, +	13 DLBCL 11BL	PET negative	NA	0.043	NA	NA	2-year OS 80
					PET positive	NA				2-year OS 29
	61	phase 2 trial	12, HIV, +	HL	PET negative	NA	NA	NA	75	2-year PFS 83
					PET positive	NA	NA	NA		

Statistically significant values are highlighted in bold.

1. AUC, area under the curve; NA, not available; 2. HL, Hodgkin's lymphoma; DLBCL, diffuse large B-cell lymphoma; BL, Burkitt's lymphoma; 3. SUV, standard uptake value; SUVmax, the maximum of SUV; SUVmean, the mean of SUV; MTV, metabolic tumor volume; TLG, total lesion glycolysis; 4. PFS, progression-free survival; OS, overall survival.

evaluate the inflammatory status at the presumed peak of the inflammatory phase. Evidence-based data has demonstrated an increased incidence of FDG PET/CT abnormalities evocative of a pulmonary infection in asymptomatic patients during the COVID-19 outbreak. However, it should be noted that FDG PET cannot substitute or integrate high-resolution CT to diagnose suspicious COVID-19 or for disease monitoring, which is different from the broader role of FDG PET in patients with HIV-associated lymphoma.

9 Conclusions and future perspectives

Taken together, PET metrics including SUV, SUL, MTV, and TLG were the most frequently observed significant parameters in the studies for differential diagnosis, response evaluation and prognostic prediction of HIV-associated lymphoma. Furthermore, baseline, interim and post-treatment PET imaging findings have been found reliable to guide clinical decision, which ultimately improves treatment outcomes and survival.

Nevertheless, the majority of current studies were single-center, retrospective studies with limited sample sizes. Additionally, the rapid development of radiomics and PET/MR will provide further lesion features and improve the PPV of PET imaging. To better understand the ability of increasingly novel radiotracers of PET imaging in HIV-associated lymphoma, including ^{68}Ga -FAPI and biomarkers of immunoPET, prospective studies with large populations and multicenter are necessary.

Author contributions

QL and TY wrote the essay. TY and YL revised the introduction and the first half of the article. XC helped to revise the framework of the article. All authors contributed to the article and approved the submitted version.

References

- Pongas GN, Ramos JC. HIV-Associated lymphomas: Progress and new challenges. *J Clin Med* (2022) 11(5):1447. doi: 10.3390/jcm11051447
- Yarchoan R, Uldrick TS. HIV-Associated cancers and related diseases. *New Engl J Med* (2018) 378(11):1029–41. doi: 10.1056/NEJMra1615896
- Robbins HA, Pfeiffer RM, Shiels MS, Li J, Hall HI, Engels EA. Excess cancers among HIV-infected people in the united states. *JNCI-Journal Natl Cancer Institute* (2015) 107(4):dju503. doi: 10.1093/jnci/dju503
- Shiels MS, Cole SR, Kirk GD, Poole C. A meta-analysis of the incidence of non-AIDS cancers in HIV-infected individuals. *J Acq Immun def synd.* (2009) 52(5):611–22. doi: 10.1097/QAI.0b013e3181b327ca
- Engels EA, Pfeiffer RM, Goedert JJ, Virgo P, McNeel TS, Scoppa SM, et al. Trends in cancer risk among people with AIDS in the united states 1980–2002. *AIDS (London England)* (2006) 20(12):1645–54. doi: 10.1097/01.aids.0000238411.75324.59
- Herida M, Mary-Krause M, Kaphan R, Cadranet J, Poizat-Martin I, Rabaud C, et al. Incidence of non-AIDS-defining cancers before and during the highly active antiretroviral therapy era in a cohort of human immunodeficiency virus-infected patients. *J Clin Oncol* (2003) 21(18):3447–53. doi: 10.1200/JCO.2003.01.096
- Lanoy E, Spano JP, Bonnet F, Guiguet M, Boué F, Cadranet J, et al. The spectrum of malignancies in HIV-infected patients in 2006 in France: the ONCOVIH study. *Int J cancer.* (2011) 129(2):467–75. doi: 10.1002/ijc.25903
- Wang C, Liu J, Liu Y. Progress in the treatment of HIV-associated lymphoma when combined with the antiretroviral therapies. *Front Oncol* (2021) 11(null):798008. doi: 10.3389/fonc.2021.798008
- Meister A, Hentrich M, Wyen C, Hübel K. Malignant lymphoma in the HIV-positive patient. *Eur J haematol* (2018) 101(1):119–26. doi: 10.1111/ejh.13082
- Biancotto A, Grivel JC, Iglehart SJ, Vanpouille C, Lisco A, Sieg SF, et al. Abnormal activation and cytokine spectra in lymph nodes of people chronically infected with HIV-1. *Blood.* (2007) 109(10):4272–9. doi: 10.1182/blood-2006-11-055764
- Liu CY, Chen BJ, Chuang SS. Primary effusion lymphoma: A timely review on the association with HIV, HHV8, and EBV. *Diagnostics (Basel Switzerland)* (2022) 12(3):713. doi: 10.3390/diagnostics12030713
- Hansen ME, Mangusan R, Lurain K, Odeny T, George J, Lu C, et al. Characteristics of patients admitted to the intensive care unit with kaposi sarcoma herpesvirus-associated diseases. *AIDS (London, England).* (2022) 36(14):1969–78. doi: 10.1097/QAD.0000000000000333
- Matinella A, Lanzafame M, Bonometti MA, Gajofatto A, Concia E, Vento S, et al. Neurological complications of HIV infection in pre-HAART and HAART era: A retrospective study. *J neurology* (2015) 262(5):1317–27. doi: 10.1007/s00415-015-7713-8
- Adams HJA, Kwee TC. An evidence-based review on the value of interim FDG-PET in assessing response to therapy in lymphoma. *Semin Oncol* (2017) 44(6):404–19. doi: 10.1053/j.seminoncol.2018.01.005
- Jiang H, Li A, Ji Z, Tian M, Zhang H. Role of radiomics-based baseline PET/CT imaging in lymphoma: Diagnosis, prognosis, and response assessment. *Mol Imaging Biol* (2022) 24(4):537–49. doi: 10.1007/s11307-022-01703-7
- Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: The lugano classification. *J Clin Oncol* (2014) 32(27):3059–68. doi: 10.1200/JCO.2013.54.8800
- Engert A, Plütschow A, Eich HT, Lohri A, Dörken B, Borchmann P, et al. Reduced treatment intensity in patients with early-stage hodgkin's lymphoma. *New Engl J Med* (2010) 363(7):640–52. doi: 10.1056/NEJMoa1000067
- André MPE, Girinsky T, Federico M, Reman O, Fortpied C, Gotti M, et al. Early positron emission tomography response-adapted treatment in stage I and II Hodgkin lymphoma: Final results of the randomized EORTC/LYSA/FIL H10 trial. *J Clin Oncol* (2017) 35(16):1786–94. doi: 10.1200/JCO.2016.68.6394
- Lang N, Crump M. PET-adapted approaches to primary therapy for advanced Hodgkin lymphoma. *Ther Adv Hematol* (2020) 11(null). doi: 10.1177/2040620720914490
- Martis N, Mounier N. Hodgkin Lymphoma in patients with HIV infection: a review. *Curr Hematologic Malignancy Rep* (2012) 7(3):228–34. doi: 10.1007/s11899-012-0125-2
- Baptista MJ, Garcia O, Morgades M, Gonzalez-Barca E, Miralles P, Lopez-Guillermo A, et al. HIV-Infection impact on clinical-biological features and outcome of diffuse large b-cell lymphoma treated with r-CHOP in the combination antiretroviral therapy era. *AIDS (London England)* (2015) 29(7):811–8. doi: 10.1097/QAD.0000000000000624
- Caponetti G, Pantanowitz L. HIV-Associated lymphadenopathy. *Ent-ear nose throat J* (2008) 87(7):374–5. doi: 10.1177/014556130808700706
- Wang Z, Zhang R, Liu L, Shen Y, Chen J, Qi T, et al. Incidence and spectrum of infections among HIV/AIDS patients with lymphoma during chemotherapy. *J Infection chemother* (2021) 27(10):1459–64. doi: 10.1016/j.jiac.2021.06.012
- Šimeková K, Nováková E, Rosolanka R, Masná J, Antolová D. Clinical course of opportunistic infections-toxoplasmosis and cytomegalovirus infection in HIV-infected patients in Slovakia. *Pathog (Basel Switzerland)* (2019) 8(4):219. doi: 10.3390/pathogens8040219
- Arora VK, Kumar SV. Pattern of opportunistic pulmonary infections in HIV sero-positive subjects: observations from pondicherry, India. *Indian J chest Dis Allied Sci* (1999) 41(3):135–44.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

26. Glushko T, He L, McNamee W, Babu AS, Simpson SA. HIV Lymphadenopathy: Differential diagnosis and important imaging features. *Am J roentgenology* (2021) 216(2):526–33. doi: 10.2214/AJR.19.22334
27. Goshen E, Davidson T, Avigdor A, Zwas TS, Levy I. PET/CT in the evaluation of lymphoma in patients with HIV-1 with suppressed viral loads. *Clin Nucl Med* (2008) 33(9):610–4. doi: 10.1097/RLU.0b013e3181813047
28. Rajeswaran G, Becker JL, Michailidis C, Pozniak AL, Padley SP. The radiology of IRIS (immune reconstitution inflammatory syndrome) in patients with mycobacterial tuberculosis and HIV co-infection: Appearances in 11 patients. *Clin radiology* (2006) 61(10):833–43. doi: 10.1016/j.crad.2006.04.007
29. Bhargava P, Chang CW, Glickman B, Sara G, Ghesani M. Persistent generalized lymphadenopathy (PGL) mimicking lymphoma on whole-body FDG PET/CT imaging. *Clin Nucl Med* (2006) 31(7):398–400. doi: 10.1097/01.rlu.0000222948.43541.0f
30. Scharko AM, Perlman SB, Pyzalski RW, Graziano FM, Sosman J, Pauza CD. Whole-body positron emission tomography in patients with HIV-1 infection. *Lancet*. (2003) 362(9388):959–61. doi: 10.1016/S0140-6736(03)14366-8
31. Liu Y. Concurrent FDG avid nasopharyngeal lesion and generalized lymphadenopathy on PET-CT imaging is indicative of lymphoma in patients with HIV infection. *AIDS Res Treat* (2012) 2012(null):764291. doi: 10.1155/2012/764291
32. Chen D, Zhu Y, Chen Y, Zhu D, Liu Z, Li T, et al. Clinical features and 18F-FDG PET/CT for distinguishing of malignant lymphoma from inflammatory lymphadenopathy in HIV-infected patients. *BMC Infect Diseases* (2022) 22(1):646. doi: 10.1186/s12879-022-07640-8
33. Mhlana JC, Durand D, Tsai HL, Durand CM, Leal JP, Wang H, et al. Differentiation of HIV-associated lymphoma from HIV-associated reactive adenopathy using quantitative FDG PET and symmetry. *Eur J Nucl Med Mol i* (2014) 41(4):596–604. doi: 10.1007/s00259-013-2671-9
34. Behin A, Hoang-Xuan K, Carpentier AF, Delattre JY. Primary brain tumours in adults. *Lancet*. (2003) 361(9354):323–31. doi: 10.1016/S0140-6736(03)12328-8
35. Marcus C, Feizi P, Hogg J, Summerfield H, Castellani R, Sriwastava S, et al. Imaging in differentiating cerebral toxoplasmosis and primary CNS lymphoma with special focus on FDG PET/CT. *Am J roentgenology* (2021) 216(1):157–64. doi: 10.2214/AJR.19.22629
36. Schroeder PC, Post MJ, Oschatz E, Stadler A, Bruce-Gregorios J, Thurnher MM. Analysis of the utility of diffusion-weighted MRI and apparent diffusion coefficient values in distinguishing central nervous system toxoplasmosis from lymphoma. *Neuroradiology*. (2006) 48(10):715–20. doi: 10.1007/s00234-006-0123-y
37. Karia SJ, McArdle DJT. AIDS-related primary CNS lymphoma. *Lancet*. (2017) 389(10085):2238. doi: 10.1016/S0140-6736(17)30056-9
38. Westwood TD, Hogan C, Julian PJ, Coutts G, Bonington S, Carrington B, et al. Utility of FDG-PETCT and magnetic resonance spectroscopy in differentiating between cerebral lymphoma and non-malignant CNS lesions in HIV-infected patients. *Eur J Radiol* (2013) 82(8):e374–9. doi: 10.1016/j.ejrad.2013.03.008
39. Hoffman JM, Waskin HA, Schifter T, Hanson MW, Gray L, Rosenfeld S, et al. FDG-PET in differentiating lymphoma from nonmalignant central nervous system lesions in patients with AIDS. *J Nucl Med* (1993) 34(4):567–75.
40. Villringer K, Jäger H, Dichgans M, Ziegler S, Poppinger J, Herz M, et al. Differential diagnosis of CNS lesions in AIDS patients by FDG-PET. *J Comput assisted tomography* (1995) 19(4):532–6. doi: 10.1097/00004728-199507000-00004
41. Lewitschnig S, Gedela K, Toby M, Kulasegaram R, Nelson M, O'Doherty M, et al. 18F-FDG PET/CT in HIV-related central nervous system pathology. *Eur J Nucl Med Mol i*. (2013) 40(9):1420–7. doi: 10.1007/s00259-013-2448-1
42. Basu S, Kung J, Houseni M, Zhuang H, Tidmarsh GF, Alavi A. Temporal profile of fluorodeoxyglucose uptake in malignant lesions and normal organs over extended time periods in patients with lung carcinoma: Implications for its utilization in assessing malignant lesions. *Q J Nucl Med Mol imaging* (2009) 53(1):9–19.
43. Kubota K, Itoh M, Ozaki K, Ono S, Tashiro M, Yamaguchi K, et al. Advantage of delayed whole-body FDG-PET imaging for tumour detection. *Eur J Nucl Med* (2001) 28(6):696–703. doi: 10.1007/s002590100537
44. Cheson BD. Role of functional imaging in the management of lymphoma. *J Clin Oncol* (2011) 29(14):1844–54. doi: 10.1200/JCO.2010.32.5225
45. Valls L, Badve C, Avril S, Herrmann K, Faulhaber P, O'Donnell J, et al. FDG-PET imaging in hematological malignancies. *Blood Rev* (2016) 30(4):317–31. doi: 10.1016/j.blre.2016.02.003
46. Network NCC. *NCCN clinical practice guidelines in oncology: B-cell lymphoma* (v.5.2022) (2022). Available at: https://www.nccn.org/professionals/physician_gls.
47. Just PA, Fieschi C, Baillet G, Galicier L, Oksenhendler E, Moretti JL. 18F-fluorodeoxyglucose positron emission tomography/computed tomography in AIDS-related burkitt lymphoma. *AIDS patient Care stds* (2008) 22(9):695–700. doi: 10.1089/apc.2008.0174
48. Hentrich M, Berger M, Wyen C, Siehl J, Rockstroh JK, Müller M, et al. Stage-adapted treatment of HIV-associated Hodgkin lymphoma: results of a prospective multicenter study. *J Clin Oncol* (2012) 30(33):4117–23. doi: 10.1200/JCO.2012.41.8137
49. Castillo JJ, Bower M, Brühlmann J, Novak U, Furrer H, Tanaka PY, et al. Prognostic factors for advanced-stage human immunodeficiency virus-associated classical Hodgkin lymphoma treated with doxorubicin, bleomycin, vinblastine, and dacarbazine plus combined antiretroviral therapy: A multi-institutional retrospective study. *Cancer*. (2015) 121(3):423–31. doi: 10.1002/cncr.29066
50. Chambers G, Frood R, Patel C, Scarsbrook A. 18F-FDG PET-CT in paediatric oncology: Established and emerging applications. *Br J Radiol* (2019) 92(1094):20180584. doi: 10.1259/bjr.20180584
51. Ngwe JY, Quek RHH, Ng DCE, Hee SW, Tao M, Lim LC, et al. High SUV uptake on FDG-PET/CT predicts for an aggressive b-cell lymphoma in a prospective study of primary FDG-PET/CT staging in lymphoma. *Ann Oncol Off J Eur Soc Med Oncol* (2009) 20(9):1543–7. doi: 10.1093/annonc/mdp030
52. Louarn N, Galicier L, Bertinchamp R, Lussato D, Montravers F, Oksenhendler É, et al. First extensive analysis of 18F-labeled fluorodeoxyglucose positron emission tomography-computed tomography in a Large cohort of patients with HIV-associated Hodgkin lymphoma: Baseline total metabolic tumor volume affects prognosis. *J Clin Oncol* (2022) 40(12):1346–55. doi: 10.1200/JCO.21.01228
53. Lawal IO, Nyakale NE, Harry LM, Modiselle MR, Ankrah AO, Msomi AP, et al. The role of f-18 FDG PET/CT in evaluating the impact of HIV infection on tumor burden and therapy outcome in patients with Hodgkin lymphoma. *Eur J Nucl Med Mol i*. (2017) 44(12):2025–33. doi: 10.1007/s00259-017-3766-5
54. Lawal IO, Ankrah AO, Popoola GO, Nyakale NE, Boshomane TG, Reyneke F, et al. 18F-FDG-PET metabolic metrics and international prognostic score for risk assessment in HIV-infected patients with Hodgkin lymphoma. *Nucl Med Commun* (2018) 39(11):1005–12. doi: 10.1097/MNM.0000000000000905
55. Reed JD, Masenge A, Buchner A, Omar F, Reyniers D, Vorster M, et al. The utility of metabolic parameters on baseline f-18 FDG PET/CT in predicting treatment response and survival in paediatric and adolescent Hodgkin lymphoma. *J Clin Med* (2021) 10(24):5979. doi: 10.3390/jcm10245979
56. Hutchings M, Loft A, Hansen M, Pedersen LM, Buhl T, Jurlander J, et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. *Blood*. (2006) 107(1):52–9. doi: 10.1182/blood-2005-06-2252
57. Gallamini A, Rigacci L, Merli F, Nassi L, Bosi A, Capodanno I, et al. The predictive value of positron emission tomography scanning performed after two courses of standard therapy on treatment outcome in advanced stage hodgkin's disease. *Haematologica-The Hematol J* (2006) 91(4):475–81.
58. Liao CC, Qin YY, Tan XH, Hu JJ, Tang Q, Rong Y, et al. Predictive value of interim PET/CT visual interpretation in the prognosis of patients with aggressive non-hodgkin's lymphoma. *OncoTargets Ther* (2017) 10(null):5727–38. doi: 10.2147/OTT.S154995
59. Okosun J, Warbey V, Shaw K, Montoto S, Fields P, Marcus R, et al. Interim fluoro-2-deoxy-D-glucose-PET predicts response and progression-free survival in patients with Hodgkin lymphoma and HIV infection. *AIDS (London England)* (2012) 26(7):861–5. doi: 10.1097/QAD.0b013e32835213b1
60. Minamimoto R, Tanuma J, Morooka M, Ito K, Okasaki M, Miyata Y, et al. Interim FDG-PET/CT as a predictor of prognosis for HIV-related malignant lymphoma: Preliminary study. *J Solid Tumors* (2013) 3(2). doi: 10.5430/jst.v3n2p1
61. Danilov AV, Li H, Press OW, Shapira I, Swinnen LJ, Noy A, et al. Feasibility of interim positron emission tomography (PET)-adapted therapy in HIV-positive patients with advanced Hodgkin lymphoma (HL): a sub-analysis of SWOG S0816 phase 2 trial. *Leukemia lymphoma* (2017) 58(2):461–5. doi: 10.1080/10428194.2016.1201573
62. Press OW, Li H, Schöder H, Straus DJ, Moskowitz CH, LeBlanc M, et al. US Intergroup trial of response-adapted therapy for stage III to IV Hodgkin lymphoma using early interim fluorodeoxyglucose-positron emission tomography imaging: Southwest oncology group S0816. *J Clin Oncol* (2016) 34(17):2020–7. doi: 10.1200/JCO.2015.63.1119
63. Engert A, Haverkamp H, Kober C, Markova J, Renner C, Ho A, et al. Reduced-intensity chemotherapy and PET-guided radiotherapy in patients with advanced stage hodgkin's lymphoma (HD15 trial): A randomised, open-label, phase 3 non-inferiority trial. *Lancet* (2012) 379(9828):1791–9. doi: 10.1016/S0140-6736(11)61940-5
64. Lamy T, Damaj G, Soubeyran P, Gyan E, Cartron G, Bouabdallah K, et al. R-CHOP 14 with or without radiotherapy in nonbulky limited-stage diffuse large b-cell lymphoma. *Blood*. (2018) 131(2):174–81. doi: 10.1182/blood-2017-07-793984
65. Persky DO, Li H, Stephens DM, Park SI, Bartlett NL, Swinnen LJ, et al. Positron emission tomography-directed therapy for patients with limited-stage diffuse large b-cell lymphoma: Results of intergroup national clinical trials network study S1001. *J Clin Oncol* (2020) 38(26):3003–11. doi: 10.1200/JCO.20.00999
66. Engert A, Kober C, Markova J. Assessment of residual bulky tumor using FDG-PET in patients with advanced-stage Hodgkin lymphoma after completion of chemotherapy: Final report of the GHSG HD15 trial. *RöFo - Fortschr auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren* (2012) 116(S 01). doi: 10.1182/blood.V116.21.764.764
67. Barrington SF, Mikhael NG, Kostakoglu L, Meignan M, Hutchings M, Müller SP, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the international conference on malignant lymphomas imaging working group. *J Clin Oncol* (2014) 32(27):3048–58. doi: 10.1200/JCO.2013.53.5229
68. Mina AA, Vakkalagadda C, Pro B. Novel therapies and approaches to Relapsed/Refractory HL beyond chemotherapy. *Cancers (Basel)* (2019) 11(3):421. doi: 10.3390/cancers11030421
69. Jerusalem G, Beguin Y, Fassotte MF, Belhocine T, Hustinx R, Rigo P, et al. Early detection of relapse by whole-body positron emission tomography in the follow-up of patients with hodgkin's disease. *Ann Oncol Off J Eur Soc Med Oncol* (2003) 14(1):123–30. doi: 10.1093/annonc/mdg011
70. El-Galaly TC, Mylam KJ, Brown P, Specht L, Christiansen I, Munksgaard L, et al. Positron emission tomography/computed tomography surveillance in patients with Hodgkin lymphoma in first remission has a low positive predictive value and high costs. *Haematologica-The Hematol J* (2012) 97(6):931–6. doi: 10.3324/haematol.2011.056010

71. Petrausch U, Samaras P, Haile SR, Veit-Haibach P, Soyka JD, Knuth A, et al. Risk-adapted FDG-PET/CT-based follow-up in patients with diffuse large b-cell lymphoma after first-line therapy. *Ann Oncol Off J Eur Soc Med Oncol* (2010) 21(8):1694–8. doi: 10.1093/annonc/mdq015

72. Annunziata S, Delgado Bolton RC, Kamani CH, Prior JO, Albano D, Bertagna F, et al. Role of 2-[18F]FDG as a radiopharmaceutical for PET/CT in patients with COVID-19: A systematic review. *Pharm (Basel Switzerland)* (2020) 13(11):377. doi: 10.3390/ph13110377



OPEN ACCESS

EDITED BY

Laura Curiel,
University of Calgary, Canada

REVIEWED BY

Partha Choudhury,
Rajiv Gandhi Cancer Institute and Research
Centre, India
Nosheen Fatima,
Aga Khan University, Pakistan

*CORRESPONDENCE

Kai Liu

✉ kliu2@houstonmethodist.org

Hong Zhao

✉ HZhao@houstonmethodist.org

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 17 October 2022

ACCEPTED 20 January 2023

PUBLISHED 02 February 2023

CITATION

Tan Y, Fang Z, Tang Y, Liu K and Zhao H
(2023) Clinical advancement of precision
theranostics in prostate cancer.
Front. Oncol. 13:1072510.
doi: 10.3389/fonc.2023.1072510

COPYRIGHT

© 2023 Tan, Fang, Tang, Liu and Zhao. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Clinical advancement of precision theranostics in prostate cancer

Yue Tan^{1,2}, Zhihui Fang^{2,3}, Yongxiang Tang², Kai Liu^{4,5*}
and Hong Zhao^{4*}

¹Hengyang Medical College, University of South China, Hengyang, Hunan, China, ²Department of Nuclear Medicine, Xiangya Hospital, Central South University, Changsha, Hunan, China, ³Key Laboratory of Biological Nanotechnology of National Health Commission, Xiangya Hospital, Central South University, Changsha, Hunan, China, ⁴Department of Systems Medicine and Bioengineering, Houston Methodist Neal Cancer Center, Weill Cornell Medicine, Houston TX, United States, ⁵Department of Gastrointestinal Surgery, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China

Theranostic approaches with positron emission tomography/computed tomography (PET/CT) or PET/magnetic resonance imaging (PET/MRI) molecular imaging probes are being implemented clinically in prostate cancer (PCa) diagnosis and imaging-guided precision surgery. This review article provides a comprehensive summary of the rapidly expanding list of molecular imaging probes in this field, including their applications in early diagnosis of primary prostate lesions; detection of lymph node, skeletal and visceral metastases in biochemical relapsed patients; and intraoperative guidance for tumor margin detection and nerve preservation. Although each imaging probe shows preferred efficacy in some applications and limitations in others, the exploration and research efforts in this field will eventually lead to improved precision theranostics of PCa.

KEYWORDS

theranostics, imaging, prostate cancer, radiotracer, PET/CT

1 Introduction

Prostate cancer (PCa) is the most frequently occurring cancer in men worldwide, with a continuously increasing incidence (1). Traditional methods for PCa diagnosis, including the digital rectal examination (DRE) and serum prostate-specific antigen (PSA) evaluation, cannot fully meet the diagnostic needs due to low accuracy and sensitivity (2). Novel methods, such as integrated positron emission tomography/computed tomography (PET/CT) or PET/magnetic resonance imaging (PET/MRI), to image ⁶⁸Gallium(⁶⁸Ga)-labeled prostate-specific membrane antigen (PSMA), which is exclusively overexpressed on clinical PCa cells, have brought great precision diagnostic capability. In addition to diagnosis, the major treatment strategy for PCa, prostatectomy, has entered the era of “precision surgery”, which requires a precise marking of the malignant tissue as intraoperative guidance. Identifying the actual position of the tumor, nerve, and lymph node has become more and more important during prostatectomy surgery. Novel intraoperative molecular imaging methods with high sensitivity, specificity, distinguishability, and safety, such as ¹¹¹In labeled PSMA, have been shown to locate the PCa lesions precisely (3); indocyanine green (ICG), a

USA Food and Drug Administration (FDA)-approved near-infrared (NIR) fluorescent agent for highlighting tissue, has been combined with ^{99m}Tc to directly and accurately recognize malignant PCa tissue and metastases to assist decision making by surgeons during operations (4). In this article, we focus on providing a comprehensive summary of all novel molecular imaging probes in PCa diagnosis and intraoperative guidance for tumor detection and nerve preservation.

2 Novel Molecular imaging methods for PCa diagnosis

2.1 Prostate-specific membrane antigen

Prostate-specific membrane antigen (PSMA) is a type II transmembrane glycoprotein encoded by the folate hydrolase 1 (FOLH1) gene. Compared with other non-specific PET tracers, PET/CT imaging targeting PSMA has important clinical value in the diagnosis and staging of PCa. PSMA is highly expressed on the surface of PCa cells and is closely correlated with tumor grade, PSA value, and prognosis. So far, two PSMA agents (^{68}Ga -PSMA11 and ^{18}F -DCFPyL) have been approved by the FDA for clinical application (5). Other PSMA tracers are also commonly used in preclinical studies and clinical trials, such as ^{68}Ga -PSMA617 and ^{18}F -PSMA-1007. A meta-analysis based on 37 studies with 4,790 patients was conducted by Perera et al. and showed that the overall sensitivity and specificity of ^{68}Ga -PSMA PET/CT for initial staging of advanced PCa were 77% and 97%, respectively (6). According to a meta-analysis by Huang, an overall pooled detection rate of 94% for ^{18}F -PSMA-1007 was demonstrated in PCa patients (7). With a combined median maximum standard uptake value (SUVmax) of 16 (3.7–77.7) for primary prostate lesions, ^{18}F -PSMA-1007 had positive predictive values of 0.90, 0.94, and 0.84 with the identification of lesions, regional lymph node metastases, and localized prostate tumors, respectively. With the comparison of regular CT imaging and bone scanning, the accuracy of PET imaging with PSMA as the target was 27% higher (92% vs 65%), as were sensitivity and specificity (85% vs 38%, 98% vs 91%) (8). Zhou et al. made a critical comparison of ^{18}F -PSMA-1007 PET/CT and ^{18}F -FDG PET/CT, which were both performed on 21 PCa patients (9). The SUVmax, mean standard uptake value (SUVmean), and tumor-to-background ratio (TBR) of ^{18}F -PSMA-1007 PET/CT were higher than those of ^{18}F -FDG PET/CT in the primary lesions and metastases, leading to a superior detection rate in the primary PCa lesions and more significant differentiations between benign lesions and metastases. In addition, the multifocality of primary PCa lesions was presented under the ^{18}F -PSMA-1007 PET/CT rather ^{18}F -FDG PET/CT, suggesting the excellent PCa lesion localization the PSMA tracer provides (Figure 1).

PSMA PET/CT imaging makes a contribution to the diagnosis of PCa metastases. A prospective single-center study demonstrated that PSMA PET/CT imaging has modest sensitivity (71.4%) and exceptional specificity (88.9%) in detecting pelvic lymph node involvement (10). In addition, PSMA PET/CT imaging combined with sentinel lymph node biopsy in primary-identified medium to high-risk PCa resulted in 94% accuracy in original lymph node staging of PCa. This cross-validation could increase the overall sensitivity of lymph node metastasis to 100%

(11). To assess the diagnostic efficacy of PSMA imaging for PCa bone metastases, a network meta-analysis that involved 45 studies with a total of 2,843 patients and 4,263 lesions was completed by Liu et al. (12). It evinced that ^{68}Ga -PSMA PET/CT had an incredible ability to visualize bone metastases, with a superiority index of 7.3, which is higher than ^{18}F -NaF, ^{11}C -choline, ^{18}F -choline, ^{18}F -fluorodeoxyglucose (FDG), and ^{18}F -fluciclovir PET/CT. Harmon et al. compared the application of ^{18}F -PSMA-PET/CT and ^{18}F -NaF in bone metastases, and 185 bone lesions were recognized by ^{18}F -NaF and/or ^{18}F -PSMA in 26 patients, in which ^{18}F -NaF significantly works better ($p < 0.001$) (13). Van Damme et al. conducted a study on 134 PCa patients, including newly diagnosed and relapsing patients, to make a comparison between PSMA imaging and whole-body magnetic resonance imaging (WB-MRI) for metastases diagnosis (14). PSMA imaging and WB-MRI were found to have no significant difference among identifying PCa patients with metastases when lymph node, skeletal, and visceral metastases were considered as a whole. However, in the subgroup of newly diagnosed PCa patients, PSMA PET/CT was better than WB-MRI for the detection of lymph node metastases, suggesting PSMA PET/CT is superior to WB-MRI for the recognition of lymph node metastasis in early PCa.

Although there are superior advantages of PSMA PET-CT imaging ascribed to the high expression of PSMA in PCa, some limitations exist. High uptake of radionuclide has been found in some benign lesions and other non-PCa malignant tumors in clinical applications (15, 16). In addition, the efficacy of PSMA PET/CT imaging is highly susceptible to the PSA level. The PSA level was positively associated with the SUV value of ^{18}F -PSMA imaging in patients with androgen deprivation therapy (ADT) (13). Combining PSA levels and treatment status, ADT patients ($n=11$) with a PSA below 2 ng/ml showed more lesions on ^{18}F -NaF than on ^{18}F -PSMA ($p=0.02$). Among patients with PSA > 2 ng/ml, ADT patients ($n=8$) showed the same or more lesions on ^{18}F -PSMA than on ^{18}F -NaF. In efficacy monitoring, ^{18}F -PSMA-1007-PET/CT has a good localization function for the biochemical recurrence (BCR) of PCa patients with a detection rate of 75% (17). Even small lymph node metastases less than 8 mm in diameter were imaged clearly. However, the remaining 25% of patients with a low level of PSA were not recognized by ^{18}F -PSMA imaging. Similar to ^{18}F -PSMA PET/CT in the relationship between the detection rate of BCR and PSA level, ^{68}Ga -PSMA PET/CT had this limitation in some studies (18, 19). Rauscher et al. analyzed subgroups of patients with very low (0.2–0.5 ng/ml) and low (0.5–1.0 ng/ml) PSA values and found that the detection rates of lesions in patients with recurrent PCa were 55% (74/134) and 74% (102/138), respectively (18). Derlin et al. investigated the imaging efficacy of the ^{68}Ga -PSMA PET/CT with more PSA level stratifications and presented much lower detection rates in patients with low PSA (< 2 ng/ml) (19). However, this limitation could be overlooked in castration-resistant prostate cancer (CRPC), which is characterized by a rising PSA. Fourquet et al. performed PSMA-PET/CT imaging in incomplete CRPC patients, which were defined as non-metastatic PCa patients after ADT treatment (20). Even for patients with PSA serum levels less than 2 ng/ml, the positive rate of PSMA PET/CT imaging could reach 70%, suggesting the high effectiveness of PSMA PET/CT imaging for CRPC-relevant diseases.

In addition, PSMA PET/CT imaging works well with the prognosis of PCa. Liu et al. used ^{68}Ga -PSMA-617 PET/CT imaging semi-quantitative analysis indicators as “imaging markers” to predict risk

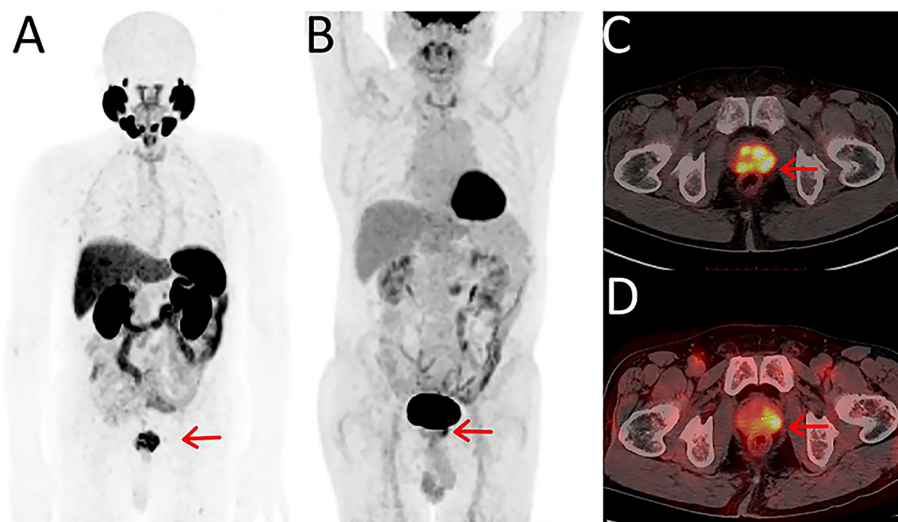


FIGURE 1
Maximum-intensity projections of PET examinations using ^{18}F -PSMA-1007 (A) and ^{18}F -FDG (B). Axial PET/CT for ^{18}F -PSMA-1007 (C) and ^{18}F -FDG (D). Reprinted with permission from Zhou et al. (9). Copyright © 2021 Zhou, Li, Jiang, Wang, Chen, Shen, You, Lu, Liao, Li and Cheng.

stratification and metastasis risk of PCa (21). Univariate logistic regression models established by SUVmax, intraprostatic PSMA-derived tumor volume (iPSMA-TV), and intraprostatic total lesion PSMA (iTL-PSMA) could be able to efficiently prewise high-risk PCa with the sensitivity and specificity of 87.5% and 50%, 62.5% and 100%, and 87.5% and 100%, respectively. A study performed by Roberts that included 848 patients after radical prostatectomy found that the SUVmax value of PSMA imaging lesions was remarkably negatively correlated with biochemical recurrence-free survival (BRFS) (22). Gleason score (GS) was also negatively correlated with BRF, and SUVmax value was an independent predictor of BRFS in patients. Roberts et al. found that increased ^{68}Ga -PSMA-11 uptake is often associated with poor pathological outcomes and provides prognostic information for progression-free survival (23). Changes in PSMA expression could be a predictive biomarker for overall survival, which may assist in personalizing therapy for PCa patients (24). ^{68}Ga -PSMA-11 PET/CT has a potential impact in guiding local lesion radiotherapy planning, which can improve the survival of castration-resistant PCa patients by adjusting the extent of radiotherapy (25). Under the guidance of PSMA PET, the mean time to PSA progression or last follow-up was 17.9 months with radiation therapy, compared with 2.9 months for patients without PSMA PET-guided local ablation radiation therapy (26). Shagera's retrospective evaluation of 37 patients with metastatic hormone-sensitive or castration-resistant prostate cancer (mHSPC or mCRPC) by testing the biochemical association between responses and different PET parameters showed that ^{68}Ga -PSMA-11 PET/CT imaging could be an effective tool for evaluating the response of metastatic PCa to taxane chemotherapy (27).

2.2 Neurotensin receptor 1

In addition to the specific molecular markers mentioned above, G protein-coupled neurotensin receptor (NTR) and its ligand neurotensin peptide (NT) have been suggested to play an important

role in PCa. Inhibiting the pathway of NTR1 has been suggested as a possible strategy to prevent the pathogenesis of this disease (28). Morgat et al. performed a pilot study of the NTR1 expression in 12 samples of normal prostate, 11 samples of benign prostatic hyperplasia (BPH), 44 samples of PCa, and 15 samples of metastatic lymph nodes (29). Compared with the negative NTR1 staining in normal prostate and BPH samples, 4 of the 44 primary tumors (9.1%) and 5 of 15 metastatic lymph nodes (33.3%) overexpressed NTR1, suggesting that NTR 1 may be a potential biomarker of PCa, especially for metastatic lymph nodes. However, the limited sample series seriously affects the reliability of this conclusion, and a larger sample size is needed in future studies.

Nevertheless, studies have suggested that NTR1 may be another molecular target that could complement PSMA imaging. Ma et al. developed novel heterodimeric probes that targeted both PSMA and NTR1 and showed significant uptake in both NTR1-positive/PSMA-negative PC-3 tumors and PSMA-positive/NTR1-negative LnCap tumors (two androgen-sensitive PCa xenografts) at the animal level (30). Zhang et al. used ^{68}Ga -DOTA-NT-20.3 animal PET imaging to scan the mice that were xenografted with PC-3, an androgen-receptor (AR)-positive, PCa cell line with no PSMA expression, suggesting that NTR1 may be a critical target for diagnosis or treatment of PCa applications with limited PSMA expression levels (31). However, the research of this tracer is still in the preclinical stage, and more preclinical and clinical studies are necessary for the exploration of its potential.

2.3 Fibroblast activation protein

First described as a cell surface antigen F19 in 1986, fibroblast activation protein (FAP) is a 760 amino acid (AA)-glycoprotein and a member of the dipeptidyl peptidase (DPP) family (32, 33). FAP shares a high AA sequence homology with DPP4, leading to its high DPP activity (34). In addition, endopeptidase activity for cleavage of

benzyloxycarbonyl-glycine-proline-7-amino-4-methylcoumarin was also found in FAP (33). Just as it initially caught people's attention for its existence in the mesenchyme of multiple cancers rather than epithelial cells, FAP was found to be overexpressed in most epithelial cancers and participate in the regulation of tumor growth and metastasis, suggesting that FAP is a potential target for tumor theranostic (35, 36). Currently, FAP inhibitors (FAPI) are mainly used for FAP-targeted PET/CT imaging (37). In 2018, Loktev et al. first reported that ^{68}Ga -FAPI-02 was used for the imaging of multiple human malignant tumors and achieved good imaging results (38). Kratochwil et al. tested the ^{68}Ga -FAPI-04 on 80 patients with 28 kinds of tumors, in which PCa patients showed intermediate uptake of ^{68}Ga -FAPI-04 with SUV of 6–12 and TBR of 3-fold (39).

Kesch et al. developed tissue microarrays (TMAs) of prostate tissues from 94 PCa patients at various stages, including primary PCa, PCa receiving ADT, CRPC, and neuroendocrine prostate cancer (NEPC), with anti-FAP antibody staining, and found the positive correlation between FAP expression and disease advancement (40). The tissues with the highest FAP expression were from CRPC patients, suggesting the potential of FAPI imaging in advanced PCa, especially CRPC. A series of case studies for the FAPI PET/CT imaging on PSMA-negative CRPC also confirmed FAPI PET/CT imaging's ability to visualize the metabolic lesions and complement the PSMA imaging (41–43). However, the issue of low sample size should be improved by large-scale clinical trials in the future.

The other weakness of FAPI PET/CT imaging is its specificity on tumor lesions. Xu et al. reported a case study of ^{68}Ga -DOTA-FAPI-04 on a PCa patient with arthritis. Compared with the prostate lesions, the arthritis site presented higher uptake of FAPI, suggesting that ^{68}Ga -DOTA-FAPI-04 may also be visualized in inflammation, possibly reducing its value in tumor diagnosis (44).

3 Additional PET agents for PCa diagnosis

3.1 ^{18}F -fluorodeoxyglucose

For tumor PET imaging, ^{18}F -fluorodeoxyglucose (^{18}F -FDG) is one of the most frequently used radiotracers. Fluorodeoxyglucose (FDG) is a glucose analogue, which is highly absorbed in tumor lesions mainly through glucose transporter-1 (GLUT1) because of its involvement in tumor cell metabolism. It has been broadly applied in clinical diagnosis, staging analysis, prognosis prediction, and treatment response monitoring of various tumors as a PET imaging agent (45). However, some patients with well-differentiated PCa had false negatives during clinical imaging (46). In addition, some benign lesions, such as inflammation, can also take up ^{18}F -FDG. Since the prostate is close to the bladder and ^{18}F -FDG is mostly egested through the urinary tract, this limits its application in the primary tumor of PCa due to the bladder urinalysis activity (9).

Although ^{18}F -FDG imaging possesses limited accuracy on primary PCa diagnosis and staging, high-grade PCa (GS = 8–10) and more aggressive metastatic PCa showed higher glycolytic activity. In a study of 148 PCa patients with biopsy GS ≥ 8 , ^{18}F -FDG PET/CT imaging detected lesions with high intraprostatic FDG uptake in 66% of patients (47). Intraprostatic FDG uptake was positively correlated

with higher pathological GS, seminal vesicle invasion, pathological lymph node metastasis, and risk of BCR, suggesting that preoperative intraprostatic FDG uptake is a composite factor for poor pathological prognostic factors. In addition, ^{18}F -FDG has a certain value in the detection of primary lesions of CRPC. Chen et al. studied 56 cases of CRPC with ^{68}Ga -PSMA and ^{18}F -FDG PET/CT examinations (48). Although overall the ^{68}Ga -PSMA is significantly better than ^{18}F -FDG PET/CT with a higher detection rate of 75.0% vs 51.8%, and more positive lesions of 135 vs 95, the incidence of patients with ^{68}Ga -PSMA-, ^{18}F -FDG+ lesions was 23.2% (13/56), which could not be ignored in the clinic. The PSA level and GS of patients with ^{68}Ga -PSMA-, ^{18}F -FDG+ lesions were higher than those of patients without ^{68}Ga -PSMA-, ^{18}F -FDG+ lesions, that 61.5% of patients with GS ≥ 8 and PSA ≥ 7.9 ng/mL carried the special lesions, suggesting that CRPC patients with high GS and PSA may take advantage of ^{18}F -FDG PET/CT imaging. ^{18}F -FDG-PET/CT is also of great value in the diagnosis of bone metastases in high-grade PCa patients (GS ≥ 8). In comparison with the bone scan, ^{18}F -FDG PET/CT is sensitive and accurate in detecting bone metastases (sensitivity: 100% vs 78.8%, specificity: 98.7% vs 98.2%) (49).

^{18}F -FDG PET imaging also has the ability to assess prognosis in PCa patients. In the study of 94 patients with primary PCa who underwent ^{18}F -FDG imaging previous to the radical prostatectomy, patients with higher SUVmax had poorer long-term survival (50). Higher intensity tracer uptake is positively associated with GLUT1 expression, stage, pathological grade, and disease progression. ^{18}F -FDG PET whole-body total lesion glycolysis (TLG) is independently associated with overall survival as a quantitative prognostic imaging biomarker in mCRPC patients receiving abiraterone or enzalutamide as first-line therapy (51). Studies have shown that SUV value and the number of lesions are also independently associated with time to hormonal therapy failure (THTF). When the sum of SUVs was divided into quartile ranges, patients in the fourth quartile had significantly lower odds of survival than patients in the first quartile. Both SUV and ^{18}F -FDG PET/CT-derived lesions provide independent prognostic information for THTF in patients with metastatic castration-sensitive PCa (52).

3.2 Choline

FDA approved the application of choline-based radiotracers (^{11}C and ^{18}F -choline) in 2012 for patients with biochemically relapsed PCa. Now both ^{11}C and ^{18}F -choline have been applied to monitor the curative effect in PCa patients. Wang et al. analyzed 46 studies and found that the combined sensitivity and specificity of ^{18}F -choline for the detection of BCR of PCa were 0.93 (95% CI, 0.85–0.98) and 0.91 (95% CI, 0.73–0.97) (53). The combined detection rate was 66%, but when PSA is in the ranges of <0.5, 0.5–0.99, 1.0–1.99, and ≥ 2 ng/mL, the detection rates were 35%, 41%, 62%, 80%, respectively. Therefore, although the choline tracer is suitable for the detection of BCR of PCa, the detection rate is not ideal when the PSA value is very low.

^{11}C and ^{18}F -choline also have implications in assessing prognosis in PCa. Jimbo et al. showed that ^{11}C -choline PET/CT assessment in mCRPC patients receiving primary docetaxel chemotherapy could predict overall treatment response and progression-free survival with blood pool-corrected SUVmax during treatment (Figure 2) (54). The

percent change in SUVmax was a significant predictor of complete response, with a greater than 20% reduction in SUVmax in 57 of 77 patients (74%), who were 3.6 times more likely to have complete remission than those patients with a reduction of SUVmax <20% after 6 cycles of primary docetaxel chemotherapy. Zhang et al. used ^{11}C -choline-PET to identify 89 patients with oligometastatic CRPC, providing a better target for stereotactic ablative radiotherapy (SABR) to improve the outcome with a median overall survival of 29.3 months (55). Garcia Vicente et al. conducted interim and end-of-treatment ^{18}F -Fluorocholine (FCH) PET/CT imaging in ^{223}Ra -treated CRPC and bone metastases patients, and the results were significantly associated with both progression-free survival and overall survival, suggesting that interim and end-of-treatment ^{18}F -FCH PET/CT imaging could be applied as predictors and even guidance during the ^{223}Ra therapy (56).

3.3 ^{18}F - Fluciclovine

^{18}F -fluciclovine (^{18}F -FACBC) was first reported by Shoup in 1999 for brain tumor imaging (57). Based on the encouraging diagnostic presentation and histologically confirmed data in patients with biochemical recurrence PCa, the FDA and European Commission (EC) approved ^{18}F -FACBC for diagnostic imaging in PCa patients with elevated PSA after pre-treatment (58), and until recently, ^{18}F -FACBC imaging has been included in the National Comprehensive Cancer Network (NCCN) guidelines for the management of BCR of PCa. A previous phase II clinical trial found the sensitivity and specificity of the scan to be 92.5% and 90.1%, respectively, for primary PCa lesions (59). Uptake of ^{18}F -FACBC was significantly increased in PCa primary lesions, and lesions with high GS (>3+4) tended to show higher uptake rates compared with low GS lesions and benign prostatic hyperplasia (60). In the diagnosis of lymph node

metastases, this study found that only 1 in 7 patients with metastatic lymph nodes showed true positive results on ^{18}F -FACBC PET/CT and PET/MRI. Another multicenter phase II study of 40 regional lymph nodes in 28 patients found that the sensitivity, specificity, diagnostic accuracy, positive predictive value, and negative predictive value of ^{18}F -FACBC imaging in lymph node analysis were 57.1% (4/7), 84.8% (28/33), 80.0% (32/40), 44.4% (4/9) and 90.3% (28/31), respectively (61). ^{18}F -FACBC PET/CT imaging has no advantage in the diagnosis of bone metastases either, possibly due to the low spatial resolution and partial volume effects caused by necrotic and mucinous components in the metastatic foci (62). A meta-analysis included 9 studies and found that the pooled sensitivity and specificity of ^{18}F -FACBC imaging of aged PCa patients (including both primary and recurrent PCa) were 86.3% and 75.9%, respectively, with a combined diagnostic odds ratio of 16.453 and heterogeneity of 30% (63). In the regional analysis, ^{18}F -FACBC-PET/CT owned a higher sensitivity and a lower specificity for the assessment of tumors in the prostate bed than in the extraprostatic region (90.4% vs 76.5%, 89% vs 45%, respectively). Filippi et al. studied the clinical data of 81 patients who underwent ^{18}F -FACBC PET/CT for BCR of PCa (64). The detection rate of ^{18}F -FACBC PET/CT in the entire cohort accounted for 76.9%, and the positive predictive value was 96.7%. This modality played an impact on the clinical management in 33 of 81 patients (40.7%), resulting in a critical amendment in treatment strategy in 30 subjects (90.9%). Like PSMA imaging, the detection rate of FACBC imaging is positively correlated with the PSA levels. When the PSA levels are in the range of 0.2-0.57, 0.58-0.99, 1-1.5 and >1.5 ng/ml, the detection rates of ^{18}F -FACBC PET/CT were 66.7%, 71.4%, 78.9% and 90, respectively. However, even at a low PSA level, ^{18}F -FACBC PET/CT imaging preserves a much higher detection rate than PSMA imaging, which is meaningful for the localization and diagnosis of lesions and has a significant impact on clinical management.

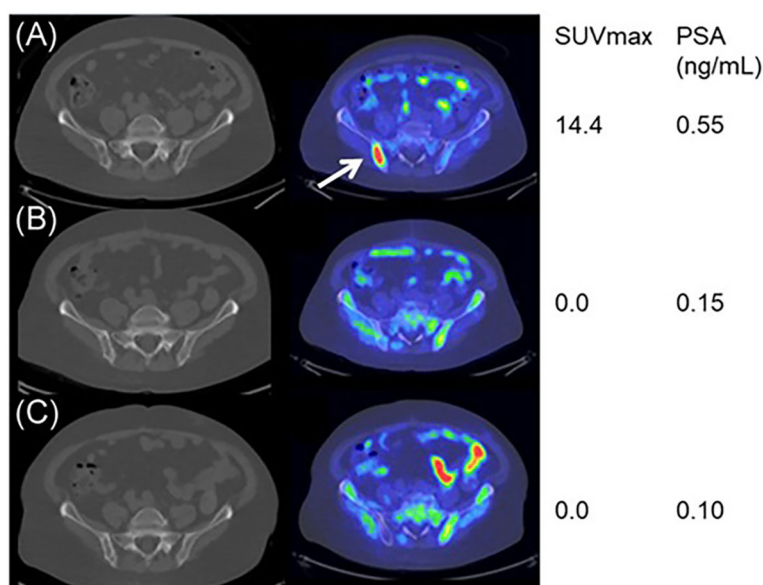


FIGURE 2
 ^{11}C -choline PET/CT imaging during the docetaxel chemotherapy for a good responder. Baseline (A), mid-course (B), and posttherapy (C) axial fused ^{11}C -choline PET/CT images demonstrating markedly choline-avid right posterior iliac bone metastasis at baseline (arrow), while nearly none at mid-course and posttherapy. Reprinted with permission from Jimbo et al. (54). Copyright ©2021 Wiley Periodicals LLC.

4 Other experimental radiotracers

4.1 Gastrin-releasing peptide receptor

Gastrin-releasing peptide receptor (GRPR) is a G protein-coupled receptor that is overexpressed in a variety of malignancies, such as breast cancer, PCa, and small cell lung cancer (65). GRPR is one of the subtypes of the bombesin (BBN) receptor, also called BB2r. As a bombesin analog, gastrin-releasing peptide (GRP) spreads over the peripheral nervous system and organs and primarily works in the gastrointestinal system through GRPR (66). As mentioned, the critical feature of GRPR is its overexpression in prostate tumor cells and underexpression in normal prostate tissue. Therefore, multiple radionuclides have been used to label bombesin analogs (GRPR agonists and antagonists), which preserve the high affinity for GRPR, to image tumors with high GRPR expressions (67, 68). At present, a variety of GRPR agonists and antagonists have emerged and been tagged with multiple radioisotopes. However, the GRPR agonists induce some gastrointestinal side effects due to the activation of GRPR. Compared with agonists, GRPR antagonists could provide better visualization with high value in the diagnosis and staging of PCa with less undesirable effects (69).

As one of the GRPR antagonists, RM26 was radiolabeled to trace the GRPR in prostate tumor tissues. In Zhang's study, both NOTA-RM26 and agonist BBN were labeled with ^{68}Ga to image the lesions in 22 PCa patients (70). The results showed that the ^{68}Ga -RM26 tracer visualized much more primary lesions and metastases with significantly higher SUVmax than ^{68}Ga -BBN PET/CT (Figure 3). Bakker et al. performed ^{68}Ga -SB3 PET/CT imaging on 10 PCa patients before radical resection with a sensitivity of 88% and a specificity of 88% in 16 lesions detected by prostatectomy pathology, suggesting that ^{68}Ga -SB3 PET/CT could be used for the detection and localization of primary PCa (71). Duan et al. compared ^{68}Ga -RM2 PET imaging with multiparametric magnetic resonance imaging (mpMRI) and ^{68}Ga -PSMA-11 PET on 41 patients with the initial diagnosis of intermediate and high-risk PCa. ^{68}Ga -RM2 and ^{68}Ga -PSMA11 had similar sensitivity and accuracy of 98%, 89% and 95%, 89%, respectively, which are significantly higher than mpMRI with 77% and 77%, for the detection of intraprostatic lesions (72). The post-prostatectomy histopathology also affirmed the ability of ^{68}Ga -RM2 PET imaging with a detection rate of 93%.

Not only for the initial diagnosis of PCa, but GRPR-targeted PET imaging could also take a role in the follow-up with the detection of BCR. Minamimoto et al. conducted a prospective study of 32 patients with BCR of PCa but negative imaging results on multiple conventional imaging modalities (CT, MRI, and $^{99\text{m}}\text{Tc}$ -MDP bone scan) (73). Among the 32 participants, 23 individuals were recognized through the ^{68}Ga -RM2 PET imaging, suggesting a detection rate of 71.8% in these patients without positive findings on conventional imaging tools. Wieser et al. also collected 16 choline-PET/CT-negative/indeterminate biochemically recurrent PCa patients to evaluate the imaging ability of ^{68}Ga -RM2-PET/CT in detecting metastatic tumors and found that tumors in 10 out of 16 patients (62.5%) could be recognized by the ^{68}Ga -RM2-PET/CT imaging (74). In addition, the expression of GRPR appears to be unassociated with PSMA, suggesting that GRPR and PSMA-targeted PET imaging could be complementary (75). Therefore, GRPR-targeted imaging could

complement other conventional modalities. Thus, GRPR tracer imaging is a promising tool for diagnosing and surveillance PCa with its high detection rate. However, the low sample size in these clinical trials critically affects the credibility of the evaluation of GRPR tracer imaging. More clinical trials with larger sample sizes are necessary for the future.

In conclusion, various imaging agents for the precision diagnosis of primary and metastatic PCa are under study, with both advantages and disadvantages (Table 1). Although there is no 100% satisfactory imaging agent for PCa at present, with the in-depth research on current imaging agents and the development of new imaging agents, multi-target combined imaging or individualized imaging may bring better clinical value to PCa patients.

5 Novel methods for intraoperative guidance of PCa precision surgery

Compared to diagnostic imaging tracers, there are fewer tracers available for guidance during PCa surgery. Herein, we present some novel intraoperative tracers, which are promising methods for PCa precision surgery in the future.

5.1 Novel methods for intraoperative tumor lesion tracing

Indocyanine green (ICG), one of the most common near-infrared (NIR) fluorophores for fluorescence-guided surgery (FGS), has been approved by the FDA for more than 60 years. It is a 776 Da, amphiphilic tricarbocyanine, water-soluble, and anionic probe. It binds to protein quickly and is confined to the intravascular compartment through intravenous injection (76). The half-life of ICG is 150-180 seconds, and it has low toxicity. Glutathione S-transferase, a transport protein, is able to make ICG through the liver and excrete into bile totally; thus, ICG can be administered repeatedly every 15 minutes during surgery to label the tissue (77). Due to its relatively low cost and widespread availability, ICG is widely used in urologic surgery, including laparoscopic and robotic adrenalectomy procedures (78, 79). In laparoscopic robot-assisted radical prostatectomy (RARP), Mangano et al. used ICG with NIR fluorescence to guide the preservation of the neurovascular bundle (80). Tobis et al. adopted ICG to highlight the renal vasculature and distinguish between normal and malignant tissue (81). Rho et al. used CT to guide the penetration of ICG through fluorescence thoracoscopy, precise location and margin resection of the radiopaque lesions were confirmed *via* C-arm fluoroscopy, and pulmonary nodules were resected with an endostapler (82). As a result, the ICG imaging guided pulmonary nodule removal was 100% in the 24 patients. However, due to the nature that ICG is a non-targeted probe with suboptimal emission characteristics for NIR-II detection, it cannot distinguish between benign and malignant tumors and can be accumulated by other tissues, which may cause false positives (83). This disadvantage was shown by Tummers et al. in a study on oncologic procedures of fluorescence-guided surgery with a high false-positive rate (62%) for the application of ICG (84).

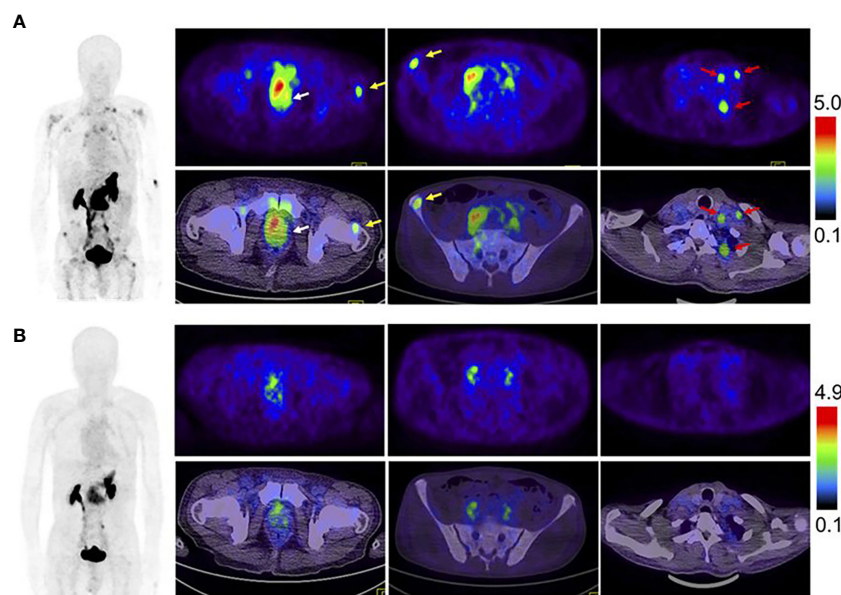


FIGURE 3
Comparison of ^{68}Ga -RM26 PET/CT (A), and ^{68}Ga -BBN PET/CT (B) in a 73-y-old man diagnosed as having PCa (white arrow) with lymph node involvement (red arrow) and bone metastasis (yellow arrow) before prostatectomy. ^{68}Ga -RM26 PET/CT detected primary tumors, multiple lymph node involvement, and bone metastasis lesion, whereas those lesions showed much lower uptake on ^{68}Ga -BBN PET/CT. Reprinted with permission from Zhang et al. (70). Copyright © 2018 by the Society of Nuclear Medicine and Molecular Imaging.

As mentioned, PSMA is a type II integral membrane glycoprotein that shows elevated expression in the majority of PCa cells (85). It is a marvelous target for image-based intraoperative guidance for accurate tumor identification due to three reasons. First, PSMA is exclusively overexpressed on tumor cells of primary PCa lesions, while its expression is consistently low in healthy prostate tissues. Second, the expression level of PSMA correlates with the Gleason grading of PCa lesions. Last, binding with the extracellular domain of PSMA normally induces internalization of the imaging agents, resulting in substantial retention of the labeling inside the tumor lesions (86). PSMA radio-guided surgery (PSMA-RGS) has been approved to be an efficient method for resecting primary tumors and metastatic lymph nodes (87).

Intravenous injection of ^{111}In -labeled PSMA-I&T to PCa patients during surgery has enabled the visualization of metastatic lymph nodes, which are normally unobtrusive and unrecognizable (88). Clinically, in patients undergoing salvage lymphadenectomy, the ^{111}In -PSMA-RGS allows intraoperative detection of small lymph node metastases with high specificity and sensitivity (89). In addition, the ^{111}In -PSMA-617 tracer also helped surgeons deal with unidentified pelvic lymph node metastases *in situ* during the surgery and resected ex vivo tissue samples to prove the successful removal (90). Except for ^{111}In -labeled PSMA ligands for detecting metastases of PCa, Robu et al. explored another ligand named $^{99\text{m}}\text{Tc}$ -mas3-y-nal-k(Sub-KuE) for PCa imaging (91). Clinically, $^{99\text{m}}\text{Tc}$ is preferable to ^{111}In , as it provides low-energy gamma rays that are more suitable for RGS due to the high sensitivity of gamma probes for collimation, and $^{99\text{m}}\text{Tc}$ has a much shorter half-life (6 hours) than ^{111}In (2.8 days), resulting in faster pharmacokinetics and lower radiation exposure for both patients and nuclear medical professionals (92).

The hybrid tracer ICG- $^{99\text{m}}\text{Tc}$ -nanocolloid combining fluorescent dye ICG with the radioactive $^{99\text{m}}\text{Tc}$ -nanocolloid, not only offers preoperative sentinel node (SN) mapping, but also provides better optical surgical guidance (93). The tracer shows no leakage into the surgical field and provides a depth estimation (>0.5–1 cm) of the nodal location, which helps to prevent surgery-related side effects (94). Another study also approved the value of the hybrid tracer in the surgical identification of lymph nodes (95). Overall, one obvious advantage of the ICG- $^{99\text{m}}\text{Tc}$ -nanocolloid tracer is that it can enable visualization of any tumor lesion or SN in their anatomical context during surgery, and its application is independent of the order of resection (primary tumor or metastasis) or the surgical setting (open or laparoscopic) (94).

5.2 Novel methods for nerve protection in PCa surgery

Iatrogenic nerve injuries are common in prostatectomy, 20% of postoperative patients suffer from urinary incontinence, and many patients experience erectile dysfunction, which can only be partially mitigated by existing nerve-sparing surgical techniques. It is challenging to intraoperatively identify the specific location of buried small peripheral nerves (PNs), but the endeavor to find new ways to protect PNs is significant (96, 97). To meet the clinical need, an ideal method for imaging PNs during the intraoperative procedure should possess the following features. First, a high specificity and a good signal-to-noise ratio are essential. Second, real-time and long-term imaging is vital for PNs to be recognized and retained during

TABLE 1 Pros and cons of PET imaging agents for prostate cancer.

Tracers	Applications	Advantages	Limitations
PSMA	Diagnosis and staging	1. Excellent TBR; 2. High sensitivity and specificity to the primary tumor and lymph node metastases; 3. Combined with MRI has a good advantage in identifying CRPC; 4. FDA approved.	1. The detection of lesions is affected by the GS; 2. Imaging of non-metastatic CRPC patients is greatly affected by PSA level; 3. Sensitivity for the diagnosis of bone metastases is affected by PSA; 4. Positive on some benign lesions and non-PCa tumors.
	Curative effect monitoring and prognosis evaluation	1. A good localization function on recurrent PCa; 2. A valuable predictive biomarker for risk stratification and metastasis; 3. Be good at predicting overall survival and BRFS; 4. An excellent guidance for local lesion radiotherapy plan; 5. Works well in chemotherapy evaluation.	The detection rate for patients with BCR is affected by PSA level and GS.
NTR1	Preclinical studies	1. Clearly expressed in NEPC with low PSMA expression, that could be another molecular target that may complement PSMA imaging; 2. Works well on lymph node metastases.	A start-up tracer only for preclinical studies.
FAPI	Diagnosis and staging	1. Independent from blood glucose level, no need for prior rest, and fast image acquisition; 2. FAP imaging has the potential to guide the treatment of mCRPC; 3. Highly positive in patients with advanced CRPC	False positives due to the uptake of FAPI in benign lesions.
FDG	Diagnosis and staging	1. Wide range of clinical applications and easy to access; 2. High detection rate for CRPC.	1. False negatives with well-differentiated PCa; 2. False positive on some benign tissues; 3. Short half-life; 4. Low efficiency in primary PCa diagnosis; 5. Low sensitivity for lymph node metastases.
	Curative effect monitoring and prognosis evaluation	Preoperative intraprostatic FDG uptake is a good prognostic factor for poor pathology	
Choline	Diagnosis and staging	1. High specificity for the diagnosis of lymph node metastases; 2. More specific to bone metastases than bone scans.	1. Much affected by neoadjuvant ADT; 2. Low sensitivity to lymph node metastases (but better than traditional MRI); 3. Poor imaging of osteogenic bone metastases; 4. Less sensitive to bone metastases than bone scans
	Curative effect monitoring and prognosis evaluation	Well prognose mCRPC patients during chemotherapy or ^{223}Ra treatment.	Not suitable for localization of lymph node metastases in patients with BCR, especially in the setting of low PSA values.
FACBC	Diagnosis and staging	1. The diagnosis of lymph node metastases is superior to choline PET imaging; 2. The overall sensitivity is higher in the prostate bed than in the extraprostatic region.	1. Limited role in local staging of PCa 2. False positive due to being non-specifically absorbed by benign prostatitis tissue; 3. The uptake was affected by GS; 4. Not suitable for bone metastases.
	Curative effect monitoring and prognosis evaluation	1. Little affected by clinicopathological parameters; 2. Works well for PCa restaging.	The localization and diagnosis of lesions in patients with BCR are affected by the level of PSA
GRPR	Diagnosis and staging	1. High detection sensitivity, specificity, and accuracy; 2. Independent of PSMA expression.	1. Low sensitivity to lymph node metastases; 2. Poor detection of bone metastases.
	Curative effect monitoring and prognosis evaluation	Lesion detection in patients with BCR of PCa is better than traditional imaging and choline PET/CT.	Affected by PSA growth rate.

surgery (98). Third, the imaging probes should have good biosafety. Last, the cost should be low enough for clinical use (97). Neurovascular dyes such as ICG and fluorescein have been used to highlight PNs in clinical settings (96). It has been shown that fluorescein was applied to visualize abnormal peroneal nerves in ganglion cyst excision procedures (99). Recently, ICG has been used to help protect critical functional structures in prostatectomy by enabling the identification of all neurovascular bundles without

increasing the operative time or complications (Figure 4) (80). These promising data indicate that iatrogenic injury can be prevented, and the operative time can be shortened with the help of fluorescence-guided imaging. According to the clinical study performed by Jin et al., in patients with ICG injected 24 hours prior to surgery, the pelvic autonomic nerves can be intraoperatively seen clearly under a NIR ray (Figure 5) (100). Due to the ubiquity of such fluorophores, it is foreseeable that surgeons will attempt the



FIGURE 4
Prostate vascularization and neurovascular bundles by ICG. Reprinted with permission from Mangano et al. (80). Copyright © 2017 Wichtig Publishing.

fluorescent nerve-targeting agents more frequently in their clinical practice (96). However, the agents can have light penetrance through the tissue of greater than 5–6 mm. Such a deep penetration causes increased light scatter, thus obscuring the specific location of PNs. In addition, ICG is not a targeted dye, and it is not able to distinguish the nerve bundle from other tissues. For example, in the surgery for deep endometriosis, the ischemic lesion, the hypogastric nerve, the pelvic plexus, and the ureter were all dyed by ICG (101).

Extensive studies have reported that PCa tumor progression is favored by innervation. Magnon et al. reported that the formation of autonomic nerve fibers in the prostate gland regulates the development and dissemination of PCa (102). Therefore, biomarkers for innervation and effective visualization methods are necessary to assess nerve density in PCa. Nerve peptide 41 (NP41) has been found as a marker to highlight peripheral nerve tissue, and fluorescent-labeled NP41 can be visualized through its binding to the motor and sensory nerves in live mice (103, 104). Hingorani et al. reported that NP41 had the best nerve-to-non-nerve contrast compared to other peptides like NP38, 40, and 42, and the average nerve-to-non-nerve signal ratio increases by 17% under fluorescent imaging compared to white light (105). NP41 is considered an excellent agent for *in vivo* tracking of nerves in rodents. Since NP41 specifically targets nerves in PCa, it has the potential for visualizing nerve density and tumor innervation in PCa. The nanoprobe named propranolol-loaded-superparamagnetic iron oxide (SPIO)-NP41 nanoparticles (PSN NPs) have been used to assess the nerve density of PCa with high sensitivity and high specificity in mice (106). Since PSN NPs had an exclusive accumulation at the tumor site, benefiting the targeted delivery of propranolol, this study showed that PSN NPs inhibited PCa tumor growth by blocking the interaction between tumor cells and sympathetic nerves in the neural tumor microenvironment.

Nevertheless, existing data on applying NP41 to ex vivo human nerve tissue provided little contrast compared to muscle (105). Hence, human NP401 (HNP401), a peptide that binds to and highlights human autonomic and motor/sensory nerves, was identified for improving the labeling of human nerves, especially for the human prostate gland, suggesting its potential guidance role in the prostatectomy for PCa patients.

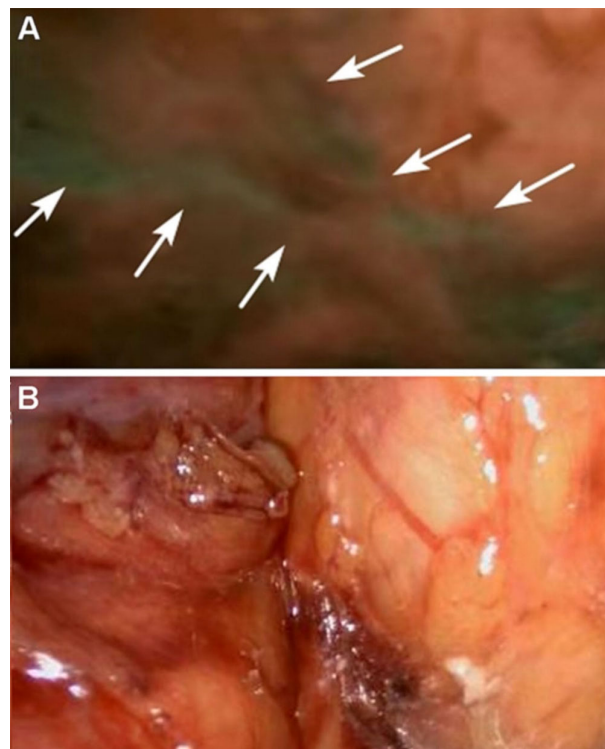


FIGURE 5
Autonomic pelvic nerves under the fluorescence (A) and under the white light (B), the sacral plexus of autonomic pelvic nerves are displayed very clearly under NIR ray (white arrows) but not clearly under white light. Reprinted with permission from Jin et al. (100). Copyright ©The Author(s), under exclusive licence to Springer Science +Business Media, LLC, part of Springer Nature 2021.

6 Expectation

Accurate and sensitive imaging using molecular probes is a promising and impactful method for early diagnosis of PCa. In addition, with molecular imaging-based intraoperative guidance, surgeons can achieve precise resection of the malignant PCa tumor as well as the metastatic lymph node, which is the trend in precision medicine. During prostatectomy, including robot-assisted radical prostatectomy (RARP), to maintain the function of the urinary system and erection postoperatively, fluorescent dye or labeled peptide hold great value in enabling visualization and protecting nerve bundles. Although each has disadvantages and limitations, all the novel methods discussed above are essential for developing early diagnosis and effective therapy of PCa. With endless exploration and research, more tracers with higher efficiency will appear to improve the precision theranostic of PCa.

Author contributions

YT, KL and HZ conceived the theme. YT, ZF and YXT conducted the writing of the manuscript. YT and ZF prepared the figures and

tables. KL and HZ edited and finalized this manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by National Natural Science Foundation of China (No: 91859207) and Houston Methodist Cornerstone Award (HZ).

Acknowledgments

The authors would like to thank Dr. Rebecca Danforth for proofreading the manuscript.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA: A Cancer J Clin* (2022) 72(1):7–33. doi: 10.3322/caac.21708
2. Mottet N, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. EAU-EANM-Estro-Esur-Siog guidelines on prostate cancer-2020 update. part 1: Screening, diagnosis, and local treatment with curative intent. *Eur Urol* (2021) 79(2):243–62. doi: 10.1016/j.eururo.2020.09.042
3. Schollhammer R, De Clermont Gallerande H, Yacoub M, Quintyn Ranty ML, Barthe N, Vimont D, et al. Comparison of the radiolabeled psma-inhibitor (111)in-Psma-617 and the radiolabeled grp-r antagonist (111)in-Rm2 in primary prostate cancer samples. *EJNMMI Res* (2019) 9(1):52. doi: 10.1186/s13550-019-0517-6
4. Cacciamani GE, Shakir A, Tafuri A, Gill K, Han J, Ahmadi N, et al. Best practices in near-infrared fluorescence imaging with indocyanine green (Nirf/Icg)-guided robotic urologic surgery: A systematic review-based expert consensus. *World J Urol* (2020) 38(4):883–96. doi: 10.1007/s00345-019-02870-z
5. Eardley I. Fda approves a second psma targeting agent for pet imaging in men with prostate cancer. *BJU Int* (2021) 128(2):127–30. doi: 10.1111/bju.15538
6. Perera M, Papa N, Roberts M, Williams M, Udovicich C, Vela I, et al. Gallium-68 prostate-specific membrane antigen positron emission tomography in advanced prostate cancer-updated diagnostic utility, sensitivity, specificity, and distribution of prostate-specific membrane antigen-avid lesions: A systematic review and meta-analysis. *Eur Urol* (2020) 77(4):403–17. doi: 10.1016/j.eururo.2019.01.049
7. Huang YT, Tseng NC, Chen YK, Huang KH, Lin HY, Huang YY, et al. The detection performance of 18 f-Prostate-Specific membrane antigen-1007 Pet/Ct in primary prostate cancer: A systemic review and meta-analysis. *Clin Nucl Med* (2022) 47(9):755–62. doi: 10.1097/rlu.0000000000004228
8. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. Prostate-specific membrane antigen pet-ct in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (PropSma): A prospective, randomised, multicentre study. *Lancet (London England)* (2020) 395(10231):1208–16. doi: 10.1016/s0140-6736(20)30314-7
9. Zhou X, Li Y, Jiang X, Wang X, Chen S, Shen T, et al. Intra-individual comparison of 18f-Psma-1007 and 18f-fdg Pet/Ct in the evaluation of patients with prostate cancer. *Front Oncol* (2020) 10:585213. doi: 10.3389/fonc.2020.585213
10. Gorin MA, Rowe SP, Patel HD, Vidal I, Mana-Ay M, Javadi MS, et al. Prostate specific membrane antigen targeted (18)F-dcfpyl positron emission Tomography/Computerized tomography for the preoperative staging of high risk prostate cancer: Results of a prospective, phase ii, single center study. *J Urol* (2018) 199(1):126–32. doi: 10.1016/j.juro.2017.07.070
11. Hinsenveld FJ, Wit EMK, van Leeuwen PJ, Brouwer OR, Donswijk ML, Tillier CN, et al. Prostate-specific membrane antigen Pet/Ct combined with sentinel node biopsy for primary lymph node staging in prostate cancer. *J Nucl Med Off publication Soc Nucl Med* (2020) 61(4):540–5. doi: 10.2967/jnumed.119.232199
12. Liu F, Dong J, Shen Y, Yun C, Wang R, Wang G, et al. Comparison of Pet/Ct and mri in the diagnosis of bone metastasis in prostate cancer patients: A network analysis of diagnostic studies. *Front Oncol* (2021) 11:736654. doi: 10.3389/fonc.2021.736654
13. Harmon SA, Bergvall E, Mena E, Shih JH, Adler S, McKinney Y, et al. A prospective comparison of (18)F-sodium fluoride Pet/Ct and psma-targeted (18)F-dcfbc Pet/Ct in metastatic prostate cancer. *J Nucl Med* (2018) 59(11):1665–71. doi: 10.2967/jnumed.117.207373
14. Van Damme J, Tombal B, Collette L, Van Nieuwenhove S, Pasoglou V, Gérard T, et al. Comparison of (68)Ga-prostate specific membrane antigen (PsmA) positron emission tomography computed tomography (Pet-ct) and whole-body magnetic

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

resonance imaging (Wb-mri) with diffusion sequences (Dwi) in the staging of advanced prostate cancer. *Cancers* (2021) 13(21):5286. doi: 10.3390/cancers13215286

15. Demirci E, Has Simsek D, Kabasakal L, Mülazimoğlu M. Abdominal splenosis mimicking peritoneal metastasis in prostate-specific membrane antigen Pet/Ct, confirmed with selective spleen Spect/Ct. *Clin Nucl Med* (2017) 42(12):e504–e5. doi: 10.1097/rlu.0000000000001878

16. Plichta KA, Graves SA, Buatti JM. Prostate-specific membrane antigen (PsmA) theranostics for treatment of oligometastatic prostate cancer. *Int J Mol Sci* (2021) 22(22):12095. doi: 10.3390/ijms222212095

17. Giesel FL, Will L, Kesch C, Freitag M, Kremer C, Merkle J, et al. Biochemical recurrence of prostate cancer: Initial results with [^{18}F]-PsmA-1007 Pet/Ct. *J Nucl Med* (2018) 59(4):632. doi: 10.2967/jnumed.117.196329

18. Rauscher I, Diwel C, Haller B, Rischpler C, Heck MM, Gschwend JE, et al. Efficacy, predictive factors, and prediction nomograms for (68)Ga-labeled prostate-specific membrane antigen-ligand positron-emission Tomography/Computed tomography in early biochemical recurrent prostate cancer after radical prostatectomy. *Eur Urol* (2018) 73(5):656–61. doi: 10.1016/j.eururo.2018.01.006

19. Derlin T, Schmuck S, Juhl C, Zörgiebel J, Schneefeld SM, Walte ACA, et al. Psa-stratified detection rates for [(68)Ga]Thp-psma, a novel probe for rapid kit-based (68)Ga-labeling and pet imaging, in patients with biochemical recurrence after primary therapy for prostate cancer. *Eur J Nucl Med Mol Imaging* (2018) 45(6):913–22. doi: 10.1007/s00259-017-3924-9

20. Fourquet A, Aveline C, Cussenot O, Créhange G, Montravers F, Talbot JN, et al. (68)Ga-Psma-11 Pet/Ct in restaging castration-resistant nonmetastatic prostate cancer: Detection rate, impact on patients' disease management and adequacy of impact. *Sci Rep* (2020) 10(1):2104. doi: 10.1038/s41598-020-58975-8

21. Liu C, Liu T, Zhang N, Liu Y, Li N, Du P, et al. (68)Ga-Psma-617 Pet/Ct: A promising new technique for predicting risk stratification and metastatic risk of prostate cancer patients. *Eur J Nucl Med Mol Imaging* (2018) 45(11):1852–61. doi: 10.1007/s00259-018-4037-9

22. Roberts MJ, Morton A, Papa N, Franklin A, Raveenthiran S, Yaxley WJ, et al. Primary tumour psma intensity is an independent prognostic biomarker for biochemical recurrence-free survival following radical prostatectomy. *Eur J Nucl Med Mol Imaging* (2022) 49(9):3289–94. doi: 10.1007/s00259-022-05756-2

23. Roberts MJ, Morton A, Donato P, Kyle S, Pattison DA, Thomas P, et al. (68)Ga-psma Pet/Ct tumour intensity pre-operatively predicts adverse pathological outcomes and progression-free survival in localised prostate cancer. *Eur J Nucl Med Mol Imaging* (2021) 48(2):477–82. doi: 10.1007/s00259-020-04944-2

24. Vlachostergios PJ, Niaz MJ, Sun M, Mosallaie SA, Thomas C, Christos PJ, et al. Prostate-specific membrane antigen uptake and survival in metastatic castration-resistant prostate cancer. *Front Oncol* (2021) 11:630589. doi: 10.3389/fonc.2021.630589

25. Calais J, Kishan AU, Cao M, Fendler WP, Eiber M, Herrmann K, et al. Potential impact of (68)Ga-Psma-11 Pet/Ct on the planning of definitive radiation therapy for prostate cancer. *J Nucl Med* (2018) 59(11):1714–21. doi: 10.2967/jnumed.118.209387

26. Lohaus F, Zöphel K, Löck S, Wirth M, Kotzerke J, Krause M, et al. Can local ablative radiotherapy revert castration-resistant prostate cancer to an earlier stage of disease? *Eur Urol* (2019) 75(4):548–51. doi: 10.1016/j.eururo.2018.11.050

27. Shagera QA, Artigas C, Karfis I, Critchi G, Chanza NM, Sideris S, et al. (68)Ga-psma Pet/Ct for response assessment and outcome prediction in metastatic prostate cancer patients treated with taxane-based chemotherapy. *J Nucl Med* (2022) 63(8):1191–8. doi: 10.2967/jnumed.121.263006

28. Zhu S, Tian H, Niu X, Wang J, Li X, Jiang N, et al. Neurotensin and its receptors mediate neuroendocrine transdifferentiation in prostate cancer. *Oncogene* (2019) 38 (24):4875–84. doi: 10.1038/s41388-019-0750-5
29. Morgat C, Chastel A, Molinier V, Schollhammer R, Macgrogan G, Vélasco V, et al. Neurotensin receptor-1 expression in human prostate cancer: A pilot study on primary tumors and lymph node metastases. *Int J Mol Sci* (2019) 20(7):1721. doi: 10.3390/ijms20071721
30. Ma X, Wang M, Wang H, Zhang T, Wu Z, Sutton MV, et al. Development of bispecific nt-psma heterodimer for prostate cancer imaging: A potential approach to address tumor heterogeneity. *Bioconjugate Chem* (2019) 30(5):1314–22. doi: 10.1021/acs.bioconjugchem.9b00252
31. Zhang-Yin J, Provost C, Cancell-Tassin G, Rusu T, Penent M, Radulescu C, et al. A comparative study of peptide-based imaging agents [(68)Ga]Ga-Psma-11, [(68)Ga]Ga-amba, [(68)Ga]Ga-Nodaga-Rgd and [(68)Ga]Ga-Dota-Nt-20.3 in preclinical prostate tumour models. *Nucl Med Biol* (2020) 84–85:88–95. doi: 10.1016/j.nucmedbio.2020.03.005
32. Rettig WJ, Chesa PG, Beresford HR, Feickert HJ, Jennings MT, Cohen J, et al. Differential expression of cell surface antigens and glial fibrillary acidic protein in human astrocytoma subsets. *Cancer Res* (1986) 46(12 Pt 1):6406–12.
33. Aertgeerts K, Levin I, Shi L, Snell GP, Jennings A, Prasad GS, et al. Structural and kinetic analysis of the substrate specificity of human fibroblast activation protein alpha. *J Biol Chem* (2005) 280(20):19441–4. doi: 10.1074/jbc.C500092200
34. Goldstein LA, Ghersi G, Piñero-Sánchez ML, Salamone M, Yeh Y, Flessate D, et al. Molecular cloning of seprase: A serine integral membrane protease from human melanoma. *Biochim Biophys Acta* (1997) 1361(1):11–9. doi: 10.1016/s0925-4439(97)00032-x
35. Garin-Chesa P, Old LJ, Rettig WJ. Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. *Proc Natl Acad Sci United States America* (1990) 87(18):7235–9. doi: 10.1073/pnas.87.18.7235
36. Fitzgerald AA, Weiner LM. The role of fibroblast activation protein in health and malignancy. *Cancer metastasis Rev* (2020) 39(3):783–803. doi: 10.1007/s10555-020-09909-3
37. Dendl K, Koerber SA, Kratochwil C, Cardinale J, Finck R, Dabir M, et al. Fap and fapi-Pet/Ct in malignant and non-malignant diseases: A perfect symbiosis? *Cancers* (2021) 13(19):4946. doi: 10.3390/cancers13194946
38. Loktev A, Lindner T, Mier W, Debus J, Altmann A, Jäger D, et al. A tumor-imaging method targeting cancer-associated fibroblasts. *J Nucl Med* (2018) 59(9):1423–9. doi: 10.2967/jnumed.118.210435
39. Kratochwil C, Flechsig P, Lindner T, Abderrahim L, Altmann A, Mier W, et al. (68) Ga-fapi Pet/Ct: Tracer uptake in 28 different kinds of cancer. *J Nucl Med* (2019) 60 (6):801–5. doi: 10.2967/jnumed.119.227967
40. Kesck C, Yirga L, Dendl K, Handke A, Darr C, Krafft U, et al. High fibroblast-Activation-Protein expression in castration-resistant prostate cancer supports the use of fapi-molecular theranostics. *Eur J Nucl Med Mol Imaging* (2021) 49(1):385–9. doi: 10.1007/s00259-021-05423-y
41. Khreish F, Rosar F, Kratochwil C, Giesel FL, Haberkorn U, Ezziddin S. Positive fapi-Pet/Ct in a metastatic castration-resistant prostate cancer patient with psma-Negative/Fdg-Positive disease. *Eur J Nucl Med Mol Imaging* (2020) 47(8):2040–1. doi: 10.1007/s00259-019-04623-x
42. Isik EG, Has-Simsek D, Sanli O, Sanli Y, Kuyumcu S. Fibroblast activation protein-targeted pet imaging of metastatic castration-resistant prostate cancer compared with 68ga-psma and 18f-fdg Pet/Ct. *Clin Nucl Med* (2022) 47(1):e54–e5. doi: 10.1097/rln.0000000000003837
43. Aryana K, Manafi-Farid R, Amini H, Divband G, Moghadam SZ. 68ga-Fapi-46 Pet/Ct in a metastatic castration-resistant prostate cancer patient with low psma expression. *Clin Nucl Med* (2022) 47(11):972–3. doi: 10.1097/RLU.0000000000004315
44. Xu T, Zhao Y, Ding H, Cai L, Zhou Z, Song Z, et al. [(68)Ga]Ga-Dota-Fapi-04 Pet/Ct imaging in a case of prostate cancer with shoulder arthritis. *Eur J Nucl Med Mol Imaging* (2021) 48(4):1254–5. doi: 10.1007/s00259-020-05028-x
45. Borea R, Favero D, Miceli A, Donegani MI, Raffa S, Gandini A, et al. Beyond the prognostic value of 2-[(18)F]Fdg Pet/Ct in prostate cancer: A case series and literature review focusing on the diagnostic value and impact on patient management. *Diagnostics (Basel Switzerland)* (2022) 12(3):581. doi: 10.3390/diagnostics12030581
46. Jadvar H. Imaging evaluation of prostate cancer with 18f-fluorodeoxyglucose Pet/Ct: Utility and limitations. *Eur J Nucl Med Mol Imaging* (2013) 40 Suppl 1(0 1):S5–10. doi: 10.1007/s00259-013-2361-7
47. Lavallée E, Bergeron M, Buteau FA, Blouin AC, Duchesnay N, Dujardin T, et al. Increased prostate cancer glucose metabolism detected by (18)F-fluorodeoxyglucose positron emission Tomography/Computed tomography in localised Gleason 8–10 prostate cancers identifies very high-risk patients for early recurrence and resistance to castration. *Eur Urol Focus* (2019) 5(6):998–1006. doi: 10.1016/j.euf.2018.03.008
48. Chen R, Wang Y, Zhu Y, Shi Y, Xu L, Huang G, et al. The added value of (18)F-fdg Pet/Ct compared with (68)Ga-psma Pet/Ct in patients with castration-resistant prostate cancer. *J Nucl Med* (2022) 63(1):69–75. doi: 10.2967/jnumed.120.262250
49. Otis-Chapados S, Goulet CR, Dubois G, Lavallée É, Dujardin T, Fradet Y, et al. (18) F-fluorodeoxyglucose positron emission Tomography/Computed tomography (Pet/Ct) is accurate for high-grade prostate cancer bone staging when compared to bone scintigraphy. *Can Urological Assoc J = J l'Association Des urologues du Canada* (2021) 15(10):301–7. doi: 10.5489/cauj.7107
50. Meziou S, Ringette Goulet C, Hovington H, Lefebvre V, Lavallée É, Bergeron M, et al. Glut1 expression in high-risk prostate cancer: Correlation with (18)F-Fdg-Pet/Ct and clinical outcome. *Prostate Cancer Prostatic Dis* (2020) 23(3):441–8. doi: 10.1038/s41391-020-0202-x
51. Wibmer AG, Morris MJ, Gonen M, Zheng J, Hricak H, Larson S, et al. Quantification of metastatic prostate cancer whole-body tumor burden with (18)F-fdg pet parameters and associations with overall survival after first-line abiraterone or enzalutamide: A single-center retrospective cohort study. *J Nucl Med Off publication Soc Nucl Med* (2021) 62(8):1050–6. doi: 10.2967/jnumed.120.256602
52. Jadvar H, Velez EM, Desai B, Ji L, Colletti PM, Quinn DI. Prediction of time to hormonal treatment failure in metastatic castration-sensitive prostate cancer with (18)F-fdg Pet/Ct. *J Nucl Med Off publication Soc Nucl Med* (2019) 60(11):1524–30. doi: 10.2967/jnumed.118.223263
53. Wang R, Shen G, Huang M, Tian R. The diagnostic role of (18)F-choline, (18)F-fluciclovine and (18)F-psma Pet/Ct in the detection of prostate cancer with biochemical recurrence: A meta-analysis. *Front Oncol* (2021) 11:684629. doi: 10.3389/fonc.2021.684629
54. Jimbo M, Andrews JR, Ahmed ME, Dundar A, Karnes RJ, Bryce AH, et al. Prognostic role of 11c-choline Pet/Ct scan in patients with metastatic castrate resistant prostate cancer undergoing primary docetaxel chemotherapy. *Prostate* (2022) 82(1):41–8. doi: 10.1002/pros.24246
55. Zhang H, Orme JJ, Abraha F, Stish BJ, Lowe VJ, Lucien F, et al. Phase ii evaluation of stereotactic ablative radiotherapy (Sabr) and immunity in (11)C-Choline-Pet/Ct-Identified oligometastatic castration-resistant prostate cancer. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2021) 27(23):6376–83. doi: 10.1158/1078-0432.Ccr-21-2510
56. García Vicente AM, Amo-Salas M, Cassinello Espinosa J, Gómez Díaz R, Soriano Castrejón A. Interim and end-treatment (18)F-fluorocholine Pet/Ct and bone scan in prostate cancer patients treated with radium 223 dichloride. *Sci Rep* (2021) 11(1):7389. doi: 10.1038/s41598-021-86759-1
57. Shoup TM, Olson J, Hoffman JM, Votaw J, Eshima D, Eshima L, et al. Synthesis and evaluation of [18f]1-Amino-3-Fluorocyclobutane-1-Carboxylic acid to image brain tumors. *J Nucl Med Off publication Soc Nucl Med* (1999) 40(2):331–8.
58. Bach-Gansmo T, Nanni C, Nieh PT, Zanon L, Bogsrud TV, Sletten H, et al. Multisite experience of the safety, detection rate and diagnostic performance of fluciclovine ((18)F) positron emission Tomography/Computerized tomography imaging in the staging of biochemically recurrent prostate cancer. *J Urol* (2017) 197(3 Pt 1):676–83. doi: 10.1016/j.juro.2016.09.117
59. Suzuki H, Inoue Y, Fujimoto H, Yonese J, Tanabe K, Fukasawa S, et al. Diagnostic performance and safety of Nmk36 (Trans-1-Amino-3-[18f]Fluorocyclobutanecarboxylic acid)-Pet/Ct in primary prostate cancer: Multicenter phase iib clinical trial. *Japanese J Clin Oncol* (2016) 46(2):152–62. doi: 10.1093/jjco/hyv181
60. Jambor I, Kuisma A, Kähkönen E, Kempainen J, Merisaari H, Eskola O, et al. Prospective evaluation of (18)F-facbc Pet/Ct and Pet/Mri versus multiparametric mri in intermediate- to high-risk prostate cancer patients (Flucipro trial). *Eur J Nucl Med Mol Imaging* (2018) 45(3):355–64. doi: 10.1007/s00259-017-3875-1
61. Suzuki H, Jinnouchi S, Kaji Y, Kishida T, Kinoshita H, Yamaguchi S, et al. Diagnostic performance of 18f-fluciclovine Pet/Ct for regional lymph node metastases in patients with primary prostate cancer: A multicenter phase ii clinical trial. *Japanese J Clin Oncol* (2019) 49(9):803–11. doi: 10.1093/jjco/hyz072
62. Bin X, Yong S, Kong QF, Zhao S, Zhang GY, Wu JP, et al. Diagnostic performance of Pet/Ct using 18f-facbc in prostate cancer: A meta-analysis. *Front Oncol* (2019) 9:1438. doi: 10.3389/fonc.2019.01438
63. Laudicella R, Albano D, Alongi P, Argiroffi G, Bauckneht M, Baldari S, et al. (18)F-facbc in prostate cancer: A systematic review and meta-analysis. *Cancers* (2019) 11 (9):1348. doi: 10.3390/cancers11091348
64. Filippi L, Bagni O, Crisafulli C, Cerio I, Brunotti G, Chiaravallotti A, et al. Detection rate and clinical impact of Pet/Ct with (18)F-facbc in patients with biochemical recurrence of prostate cancer: A retrospective bicentric study. *Biomedicines* (2022) 10 (1):177. doi: 10.3390/biomedicines10010177
65. Chaudhary PK, Kim S. An insight into gpcr and G-proteins as cancer drivers. *Cells* (2021) 10(12):3288. doi: 10.3390/cells10123288
66. Huang H, Ekama G, Deng YF, Chen GH, Wu D. Identifying the mechanisms of sludge reduction in the sulfidogenic oxic-settling anaerobic (Sosa) process: Side-stream sulfidogenesis-intensified sludge decay and mainstream extended aeration. *Water Res* (2021) 189:116608. doi: 10.1016/j.watres.2020.116608
67. Biddlecombe GB, Rogers BE, de Visser M, Parry JJ, de Jong M, Erion JL, et al. Molecular imaging of gastrin-releasing peptide receptor-positive tumors in mice using 64cu- and 86y-Dota-(Pro1,Tyr4)-Bombesin(1-14). *Bioconjugate Chem* (2007) 18(3):724–30. doi: 10.1021/bc060281l
68. Chatlial KL, Franssen GM, van Weerden WM, McBride WJ, Laverman P, de Blois E, et al. Preclinical comparison of Al18f- and 68ga-labeled gastrin-releasing peptide receptor antagonists for pet imaging of prostate cancer. *J Nucl Med Off publication Soc Nucl Med* (2014) 55(12):2050–6. doi: 10.2967/jnumed.114.141143
69. Baratto L, Jadvar H, Iagaru A. Prostate cancer theranostics targeting gastrin-releasing peptide receptors. *Mol Imaging Biol* (2018) 20(4):501–9. doi: 10.1007/s11307-017-1151-1
70. Zhang J, Niu G, Fan X, Lang L, Hou G, Chen L, et al. Pet using a grpr antagonist (68)Ga-Rm26 in healthy volunteers and prostate cancer patients. *J Nucl Med* (2018) 59 (6):922–8. doi: 10.2967/jnumed.117.198929
71. Bakker IL, Fröberg AC, Busstra MB, Verzijlbergen JF, Konijnenberg M, van Leenders G, et al. Grpr antagonist (68)Ga-Sb3 Pet/Ct imaging of primary prostate

cancer in therapy-naïve patients. *J Nucl Med* (2021) 62(11):1517–23. doi: 10.2967/jnumed.120.258814

72. Duan H, Baratto L, Fan RE, Soerensen SJC, Liang T, Chung BI, et al. Correlation of (68)Ga-Rm2 pet with post-surgery histopathology findings in patients with newly diagnosed intermediate- or high-risk prostate cancer. *J Nucl Med* (2022) 63(12):1829–35. doi: 10.2967/jnumed.122.263971

73. Minamimoto R, Sonni I, Hancock S, Vasanawala S, Loening A, Gambhir SS, et al. Prospective evaluation of (68)Ga-Rm2 Pet/Mri in patients with biochemical recurrence of prostate cancer and negative findings on conventional imaging. *J Nucl Med Off publication Soc Nucl Med* (2018) 59(5):803–8. doi: 10.2967/jnumed.117.197624

74. Wieser G, Popp I, Christian Rischke H, Drendel V, Grosu AL, Bartholomä M, et al. Diagnosis of recurrent prostate cancer with Pet/Ct imaging using the gastrin-releasing peptide receptor antagonist (68)Ga-Rm2: Preliminary results in patients with negative or inconclusive [(18)F]Fluoroethylcholine-Pet/Ct. *Eur J Nucl Med Mol Imaging* (2017) 44(9):1463–72. doi: 10.1007/s00259-017-3702-8

75. Touijer KA, Michaud L, Alvarez HAV, Gopalan A, Kossatz S, Gonen M, et al. Prospective study of the radiolabeled grpr antagonist Bay86-7548 for positron emission Tomography/Computed tomography imaging of newly diagnosed prostate cancer. *Eur Urol Oncol* (2019) 2(2):166–73. doi: 10.1016/j.euo.2018.08.011

76. Nagaya T, Nakamura YA, Choyke PL, Kobayashi H. Fluorescence-guided surgery. *Front Oncol* (2017) 7:314. doi: 10.3389/fonc.2017.00314

77. Kaplan-Marans E, Fulla J, Tomer N, Bilal K, Palese M. Indocyanine green (Icg) in urologic surgery. *Urology* (2019) 132:10–7. doi: 10.1016/j.urology.2019.05.008

78. Lee JYK, Pierce JT, Thawani JP, Zeh R, Nie S, Martinez-Lage M, et al. Near-infrared fluorescent image-guided surgery for intracranial meningioma. *J Neurosurg* (2018) 128(2):380–90. doi: 10.3171/2016.10.Jns161636

79. Tipirneni KE, Warram JM, Moore LS, Prince AC, de Boer E, Jani AH, et al. Oncologic procedures amenable to fluorescence-guided surgery. *Ann Surg* (2017) 266(1):36–47. doi: 10.1097/sla.0000000000002127

80. Mangano MS, De Gobbi A, Beniamin F, Lamon C, Ciaccia M, Maccatrozzo L. Robot-assisted nerve-sparing radical prostatectomy using near-infrared fluorescence technology and indocyanine green: Initial experience. *Urologia* (2018) 85(1):29–31. doi: 10.5301/uj.5000244

81. Tobis S, Knopf J, Silvers C, Yao J, Rashid H, Wu G, et al. Near infrared fluorescence imaging with robotic assisted laparoscopic partial nephrectomy: Initial clinical experience for renal cortical tumors. *J Urol* (2011) 186(1):47–52. doi: 10.1016/j.juro.2011.02.2701

82. Rho J, Lee JW, Quan YH, Choi BH, Shin BK, Han KN, et al. Fluorescent and iodized emulsion for preoperative localization of pulmonary nodules. *Ann Surg* (2021) 273(5):989–96. doi: 10.1097/sla.0000000000003300

83. Hu Z, Fang C, Li B, Zhang Z, Cao C, Cai M, et al. First-in-Human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-Infrared-I/Ii windows. *Nat Biomed Eng* (2020) 4(3):259–71. doi: 10.1038/s41551-019-0494-0

84. Tummers QR, Hoogstins CE, Peters AA, de Kroon CD, Trimbois JB, van de Velde CJ, et al. The value of intraoperative near-infrared fluorescence imaging based on enhanced permeability and retention of indocyanine green: Feasibility and false-positives in ovarian cancer. *PLoS One* (2015) 10(6):e0129766. doi: 10.1371/journal.pone.0129766

85. Lütje S, Slavik R, Fendler W, Herrmann K, Eiber M. PsmA ligands in prostate cancer - probe optimization and theranostic applications. *Methods (San Diego Calif)* (2017) 130:42–50. doi: 10.1016/j.jymeth.2017.06.026

86. Derks YHW, Löwik D, Sedelaar JPM, Gotthardt M, Boerman OC, Rijpkema M, et al. PsmA-targeting agents for radio- and fluorescence-guided prostate cancer surgery. *Theranostics* (2019) 9(23):6824–39. doi: 10.7150/thno.36739

87. Maurer T, Robu S, Schottelius M, Schwamborn K, Rauscher I, van den Berg NS, et al. (99m)Technetium-based prostate-specific membrane antigen-radioguided surgery in recurrent prostate cancer. *Eur Urol* (2019) 75(4):659–66. doi: 10.1016/j.eururo.2018.03.013

88. Maurer T, Weirich G, Schottelius M, Weineisen M, Frisch B, Okur A, et al. Prostate-specific membrane antigen-radioguided surgery for metastatic lymph nodes in prostate cancer. *Eur Urol* (2015) 68(3):530–4. doi: 10.1016/j.eururo.2015.04.034

89. Rauscher I, Düwel C, Wirtz M, Schottelius M, Wester HJ, Schwamborn K, et al. Value of (111) in-Prostate-Specific membrane antigen (PsmA)-radioguided surgery for

salvage lymphadenectomy in recurrent prostate cancer: Correlation with histopathology and clinical follow-up. *BJU Int* (2017) 120(1):40–7. doi: 10.1111/bju.13713

90. Mix M, Reichel K, Stoykow C, Bartholomä M, Drendel V, Gourni E, et al. Performance of (111)in-labelled psmA ligand in patients with nodal metastatic prostate cancer: Correlation between tracer uptake and histopathology from lymphadenectomy. *Eur J Nucl Med Mol Imaging* (2018) 45(12):2062–70. doi: 10.1007/s00259-018-4094-0

91. Robu S, Schottelius M, Eiber M, Maurer T, Gschwend J, Schwaiger M, et al. Preclinical evaluation and first patient application of 99mtc-PsmA-I&S for spect imaging and radioguided surgery in prostate cancer. *J Nucl Med* (2017) 58(2):235–42. doi: 10.2967/jnumed.116.178939

92. Povoski SP, Neff RL, Mojzisek CM, O'Malley DM, Hinkle GH, Hall NC, et al. A comprehensive overview of radioguided surgery using gamma detection probe technology. *World J Surg Oncol* (2009) 7:11. doi: 10.1186/1477-7819-7-11

93. van der Poel HG, Buckle T, Brouwer OR, Valdés Olmos RA, van Leeuwen FW. Intraoperative laparoscopic fluorescence guidance to the sentinel lymph node in prostate cancer patients: Clinical proof of concept of an integrated functional imaging approach using a multimodal tracer. *Eur Urol* (2011) 60(4):826–33. doi: 10.1016/j.eururo.2011.03.024

94. KleinJan GH, van Werkhoven E, van den Berg NS, Karakullukcu MB, Zijlman H, van der Hage JA, et al. The best of both worlds: A hybrid approach for optimal pre- and intraoperative identification of sentinel lymph nodes. *Eur J Nucl Med Mol Imaging* (2018) 45(11):1915–25. doi: 10.1007/s00259-018-4028-x

95. Meershoek P, Buckle T, van Oosterom MN, KleinJan GH, van der Poel HG, van Leeuwen FWB. Can intraoperative fluorescence imaging identify all lesions while the road map created by preoperative nuclear imaging is masked? *J Nucl Med* (2020) 61(6):834–41. doi: 10.2967/jnumed.119.235234

96. Walsh EM, Cole D, Tipirneni KE, Bland KI, Udayakumar N, Kasten BB, et al. Fluorescence imaging of nerves during surgery. *Ann Surg* (2019) 270(1):69–76. doi: 10.1097/sla.0000000000003130

97. Xu H, Chen J, Feng Z, Fu K, Qiao Y, Zhang Z, et al. Shortwave infrared fluorescence in vivo imaging of nerves for minimizing the risk of intraoperative nerve injury. *Nanoscale* (2019) 11(42):19736–41. doi: 10.1039/c9nr06066a

98. Gibbs-Strauss SL, Nasr KA, Fish KM, Khullar O, Ashitate Y, Siclovian TM, et al. Nerve-highlighting fluorescent contrast agents for image-guided surgery. *Mol Imaging* (2011) 10(2):91–101. doi: 10.2310/7290.2010.00026

99. Stone JJ, Graffeo CS, de Ruiter GCW, Rock MG, Spinner RJ. Intraoperative intravenous fluorescein as an adjunct during surgery for peroneal intraneural ganglion cysts. *Acta neurochirurgica* (2018) 160(3):651–4. doi: 10.1007/s00701-018-3477-0

100. Jin H, Zheng L, Lu L, Cui M. Near-infrared intraoperative imaging of pelvic autonomic nerves: A pilot study. *Surg endoscopy* (2022) 36(4):2349–56. doi: 10.1007/s00464-021-08512-z

101. Kanno K, Aiko K, Yanai S, Sawada M, Sakate S, Andou M. Clinical use of indocyanine green during nerve-sparing surgery for deep endometriosis. *Fertility sterility* (2021) 116(1):269–71. doi: 10.1016/j.fertnstert.2021.03.014

102. Magnon C, Hall SJ, Lin J, Xue X, Gerber L, Freedland SJ, et al. Autonomic nerve development contributes to prostate cancer progression. *Sci (New York NY)* (2013) 341(6142):1236361. doi: 10.1126/science.1236361

103. Hussain T, Mastrodimos MB, Raju SC, Glasgow HL, Whitney M, Friedman B, et al. Fluorescently labeled peptide increases identification of degenerated facial nerve branches during surgery and improves functional outcome. *PLoS One* (2015) 10(3):e0119600. doi: 10.1371/journal.pone.0119600

104. Whitney MA, Crisp JL, Nguyen LT, Friedman B, Gross LA, Steinbach P, et al. Fluorescent peptides highlight peripheral nerves during surgery in mice. *Nat Biotechnol* (2011) 29(4):352–6. doi: 10.1038/nbt.1764

105. Hingorani DV, Whitney MA, Friedman B, Kwon JK, Crisp JL, Xiong Q, et al. Nerve-targeted probes for fluorescence-guided intraoperative imaging. *Theranostics* (2018) 8(15):4226–37. doi: 10.7150/thno.23084

106. You H, Shang W, Min X, Weinreb J, Li Q, Leapman M, et al. Sight and switch off: Nerve density visualization for interventions targeting nerves in prostate cancer. *Sci Adv* (2020) 6(6):eaax6040. doi: 10.1126/sciadv.aax6040

Glossary

AA	Amino acid
ADT	Androgen deprivation therapy
AR	Androgen-receptor
BCR	Biochemical recurrence
BPH	Benign prostatic hyperplasia
BRFS	Biochemical recurrence-free survival
CRPC	Castration-resistant prostate cancer
DPP	Dipeptidyl peptidase
DRE	Digital rectal examination
FACBC	Fluciclovine
FAP	Fibroblast activation protein
FAPI	FAP inhibitors
FCH	Fluorocholine
FDA	Food and Drug Administration
FDG	Fluorodeoxyglucose
FGS	Fluorescence-guided surgery
FOLH1	Folate hydrolase 1
Ga	Gallium
GLUT1	Glucose transporter-1
GS	Gleason score
GRPR	Gastrin-releasing peptide receptor
HNP401	Human NP401
ICG	Indocyanine green
iPSMA-TV	Intraprostatic PSMA-derived tumor volume
iTL-PSMA	Intraprostatic total lesion PSMA
mHSPC or mCRPC	Metastatic hormone-sensitive or castration-resistant prostate cancer
mpMRI	Multiparametric magnetic resonance imaging
SUV _{max}	Maximum standard uptake value
SUV _{mean}	Mean standard uptake value
NCCN	National Comprehensive Cancer Network
NEPC	Neuroendocrine prostate cancer
NIR	Near-infrared
NP41	Nerve peptide 41
NPs	Nanoparticles
NT	Neurotensin peptide
NTR	Neurotensin receptor
PCa	Prostate cancer
PET/CT	Positron emission tomography/computed tomography
PET/MRI	PET/magnetic resonance imaging

(Continued)

Continued

PNs	Peripheral nerves
PSA	Prostate-specific antigen
PSMA	Prostate-specific membrane antigen
PSMA-RGS	PSMA radio-guided surgery
RARP	Robot-assisted radical prostatectomy
SABR	Stereotactic ablative radiotherapy
SN	Sentinel node
SPIO	Superparamagnetic iron oxide
TBR	Tumor-to-background ratio
TLG	Total lesion glycolysis
THTF	Time to hormonal therapy failure
TMA	Tissue microarrays
WB-MRI	Whole-body magnetic resonance imaging



OPEN ACCESS

EDITED BY

Laura Curiel,
University of Calgary, Canada

REVIEWED BY

Bilgin Kadri Aribas,
Bülent Ecevit University, Türkiye
Yixing Yu,
The First Affiliated Hospital of Soochow
University, China

*CORRESPONDENCE

Wengui Xu
✉ shanzibing@126.com

[†]These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 06 November 2022

ACCEPTED 24 January 2023

PUBLISHED 08 February 2023

CITATION

Cao L, Zhang L and Xu W (2023) Small hyperattenuating adrenal nodules in patients with lung cancer: Differentiation of metastases from adenomas on biphasic contrast-enhanced computed tomography.
Front. Oncol. 13:1091102.
doi: 10.3389/fonc.2023.1091102

COPYRIGHT

© 2023 Cao, Zhang and Xu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Small hyperattenuating adrenal nodules in patients with lung cancer: Differentiation of metastases from adenomas on biphasic contrast-enhanced computed tomography

Lixiu Cao^{1,2†}, Libo Zhang^{1†} and Wengui Xu^{1*}

¹Department of Molecular Imaging and Nuclear Medicine, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China, ²Department of Emission Computed Tomography, Tangshan People's Hospital, Tangshan, Hebei, China

Objective: The objective of this study was to evaluate the value of biphasic contrast-enhanced computed tomography (CECT) in the differential diagnosis of metastasis and lipid-poor adenomas (LPAs) in lung cancer patients with unilateral small hyperattenuating adrenal nodule.

Materials and methods: This retrospective study included 241 lung cancer patients with unilateral small hyperattenuating adrenal nodule (metastases, 123; LPAs, 118). All patients underwent plain chest or abdominal computed tomography (CT) scan and biphasic CECT scan, including arterial and venous phases. Qualitative and quantitative clinical and radiological characteristics of the two groups were compared using univariate analysis. An original diagnostic model was developed using multivariable logistic regression, and then, according to odds ratio (OR) of the risk factors of metastases, a diagnostic scoring model was developed. The areas under the receiver operating characteristic curves (AUCs) of the two diagnostic models were compared by DeLong test.

Results: Compared with LPAs, metastases were older and showed more frequently irregular in shape and cystic degeneration/necrosis (all $p < 0.05$). Enhancement ratios on venous (ERV) and arterial (ERA) phase of LPAs were noticeably higher than that of metastases, whereas CT values in unenhanced phase (UP) of LPAs were noticeably lower than that of metastases (all $p < 0.05$). Compared with LPAs, the proportions of male and III/IV clinical stage and small-cell lung cancer (SCLL) were significantly higher for metastases (all $p < 0.05$). As for peak enhancement phase, LPAs showed relatively faster wash-in and earlier wash-out enhancement pattern than metastases ($p < 0.001$). Multivariate analysis revealed age ≥ 59.5 years (OR: 2.269; $p = 0.04$), male (OR: 3.511; $p = 0.002$), CT values in UP ≥ 27.5 HU (OR: 6.968; $p < 0.001$), cystic degeneration/necrosis (OR: 3.076; $p = 0.031$), ERV ≤ 1.44 (OR: 4.835; $p < 0.001$), venous phase or equally enhanced (OR: 16.907; $p < 0.001$ or OR: 14.036; $p < 0.001$), and clinical stage II or III or IV (OR: 3.550; $p = 0.208$ or OR: 17.535; $p = 0.002$ or OR: 20.241; $p = 0.001$) were risk factors for diagnosis of

metastases. AUCs of the original diagnostic model and the diagnostic scoring model for metastases were 0.919 (0.883–0.955) and 0.914 (0.880–0.948), respectively. There was no statistical significance of AUC between the two diagnostic model ($p = 0.644$).

Conclusions: Biphasic CECT performed well diagnostic ability in differentiating metastases from LAPs. The diagnostic scoring model is easy to popularize due to simplicity and convenience.

KEYWORDS

small hyperattenuating adrenal nodules, metastases of lung cancer, lipid-poor adenomas, biphasic contrast enhanced CT, diagnostic scoring model

1 Introduction

Worldwide, the incidence and mortality rates of lung cancer rank first for men and second for women (1). The high mortality rate may be associated with the development of metastasis. A frequent site of metastatic spread is the adrenal glands (2). The most common malignant tumor involving the adrenal gland is metastasis, which is also the second most common adrenal tumor after adenoma. Previous studies have showed that lung cancer is the most common primary cancer of adrenal metastases (3–5); meanwhile, approximately half of adrenal tumors in patients with lung cancer were metastases (6). The qualitative diagnosis of adrenal lesions in patients with lung cancer is therefore critical to stage, direct therapy, and predict prognosis of lung cancer.

Based on endocrine function tests, clinical symptoms, and radiologic characteristics, a specific diagnosis of adrenal nodules can be achieved in many patients with lung cancer. However, when patients have nonfunctioning and indeterminate (unilateral small [long diameter, LD] ≤ 3 cm) hyperattenuating [CT values in UP >10 HU]) adrenal nodules based on conventional chest or abdominal biphasic CECT, making a correct diagnosis of metastases immediately without additional diagnostic steps is challenging because the imaging features of metastases overlap with those of LAPs (7–9), especially in patients with lung cancer, and then additional confirmatory steps, such as adrenal washout CT, MRI, PET/CT, and biopsy, may be required (10–14). Although adrenal washout CT used for the characterization of adrenal nodules has relatively high sensitivity and specificity (15), it will add additional radiation and medical costs to patients, and the delay scan time is too long (10). Moreover, previous studies reported that the enhanced washout ratios of some adrenal metastases are similar to that of LAPs, leading to misdiagnosis

(16). As the most sensitive examination, chemical-shift MRI still indicates indeterminate findings in approximately 10–20% of LAPs, and not all patients have high-quality MRI images (17). For PET/CT, the (18F)-fluorodeoxyglucose uptake in adenomas and metastases has a certain overlap (18). Additionally, PET/CT is not generally utilized in most institutions, and several days of waiting often occur. To achieve accurate diagnosis of indeterminate adrenal nodules, most doctors will choose invasive diagnostic procedures, such as biopsy or surgical resection, which may result in unnecessary patient anxiety and over-diagnosis and some complications (19).

Advanced image analysis technique such as radiomics has been proved to be able to differentiate adrenal tumors, especially benign and malignant lesions. However, radiomics is not used routinely in clinical practice, because it requires computational expertise and its reliability is still uncertain. Thus, there is a need for the development of simple and non-invasive rule-in or rule-out method for effectively characterizing these indeterminate adrenal nodules in lung cancer patients who await treatment for a potentially fatal disease.

Foti et al. (20) demonstrated that most adrenal metastases had a slower wash-in characteristic from biphasic CECT than adenomas. Lee et al. (21) also proved that wash-in characteristics from unenhanced to portal phase can effectively distinguish hyperattenuating adrenal tumors in lung cancer patients as to wash-out features on adrenal CT. These studies indicate that early biphasic CECT (it is equivalent to the conventional chest or abdominal biphasic CECT) without 15-min delay scan has the ability to differentiate whether adrenal lesions are metastasis or benign. In this way, not only the scanning time can be greatly reduced by omitting the 15-min delay scan but also no specific software or hardware is required to calculate the wash-in parameters based on biphasic CECT. However, previous studies only paid close attention to individual characteristic, and other quantitative and qualitative radiological characteristics were not taken into consideration for characterizing adrenal lesions. In addition, clinical stage and histologic subtypes of lung cancer were not analyzed. Moreover, data on specific “indeterminate adrenal nodules” are still very limit. Comprehensive differential diagnostic criteria may be required for “indeterminate adrenal nodules” in the characterization of lesions as metastases of lung cancer. Therefore, the purpose of the present study was to investigate whether biphasic

Abbreviations: LAPs, lipid-poor adrenal adenomas; AUC, area under the receiver operating characteristic curve; CT, computed tomography; CECT, contrast-enhanced computed tomography; LD, long diameter; SD, short diameter; CI, confidence interval; ROI, region of interest; HU, Hounsfield units; OR, odds ratio; AP: arterial phase; VP: venous phase; UP: unenhanced phase; ERA, enhancement ratio on arterial phase; ERV, enhancement ratio on venous phase; PPV, positive predictive value; NPV, negative predictive value; ROC, receiver operating characteristic curve.

CECT could distinguish between metastases and LAPs for indeterminate adrenal nodules in lung cancer patients.

2 Materials and methods

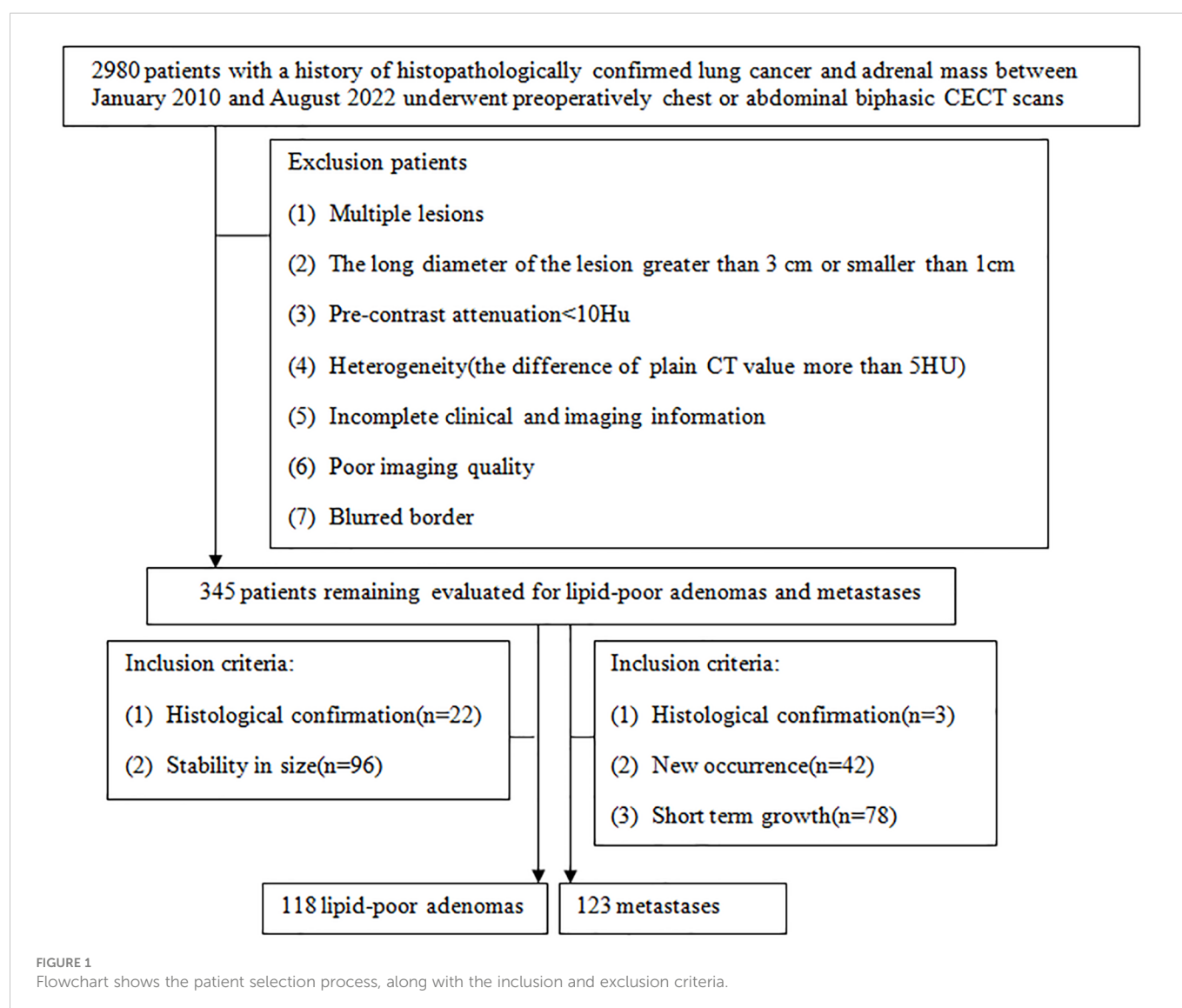
2.1 Patients

This study had been approved by Tangshan People's Hospital Institutional Ethics Committee. Patients who met the following inclusion criteria from February 2010 to August 2022 were included: (1) patients with a history of histopathologically confirmed lung cancer before or after undergoing chest or abdominal biphasic CECT; (2) indeterminate adrenal nodules: unilateral small ($1\text{ cm} \leq \text{LD} \leq 3\text{ cm}$) hyperattenuating (CT values in UP $> 10\text{ HU}$) adrenal nodules. There were two main reasons for the use of the cutoff of 1 cm for LD of adrenal tumor: (a) to increase confidence in the presence of a truly focal adrenal tumor and (b) to allow sufficient tumor volume for reliable quantitative measurement techniques; (3) complete clinical and imaging information. There were three eligibility criteria for diagnosing metastases: (1)

histologically confirmed resection specimen or needle biopsy ($n = 3$), (2) new occurrence of a lesion in the adrenal gland on follow-up CT ($n = 42$), and (3) the total sum of the nodule of the same patient grew by 20% within 6 months (22) ($n = 78$). There were two eligibility criteria for diagnosing LAPs: (1) surgically excised and histopathological assessment ($n = 22$) and (2) stability in size after at least a 1 year interval ($n = 96$; mean follow-up time = 813 days \pm 301). Finally, this study comprised 123 metastases and 118 LAPs (Figure 1).

2.2 Image protocol

Owing to retrospective analysis, this study used two CT scanners, Ingenuity core 64 (Philips Healthcare) and GE Discovery CT750 HD (GE Healthcare). All the patients undergone chest or abdominal plain and biphasic CECT scan, including arterial (approximately 30 s) and venous phase (approximately 60 s) after 80–100 ml of non-ionic contrast agent iodopamil (350 mg I/ml) was infused with a high-pressure syringe at a rate of 3.5 ml/s. The scanning parameters and image reconstruction are listed in [Supplementary Table S1](#).



2.3 Imaging analysis

The short diameter (SD), LD, right or left, shape (regular: round or oval; irregular: unrounded), cystic degeneration/necrosis (low-density region without enhancement), CT values in UP, arterial phase (AP), and venous phase (VP) of adrenal nodules were independently measured and evaluated by two radiologists with 5 and 9 years of abdominal CT diagnosis experience on thin-sliced CT images. They were both blind to pathological results and clinical information. Disagreement was settled by consensus. LD and SD should be measured on the largest cross section of the adrenal nodules. When measuring the CT value, the region of interest (ROI) should include two-thirds of the maximum axial area of the nodules, excluding adjacent fat. In addition, we should measure three times for the attenuation values and record the average value of these values as the final result. The calculation formulas of ERA and ERV were $ERA = (CT \text{ values in AP} - CT \text{ values in UP}) / CT \text{ values in UP}$ and $ERV = (CT \text{ values in VP} - CT \text{ values in UP}) / CT \text{ values in UP}$, respectively. We defined the phase in which maximum enhancement level was 5 HU greater than another phase as the peak enhancement phase, otherwise, equally enhanced when the difference of enhancement level was less than 5 HU between arterial and venous phase (23, 24).

2.4 Statistical analysis

All data were analyzed using SPSS 21, R software (version 4.2.1; <http://www.rproject.org>), and MedCalc 20.0.22. The χ^2 test or Fisher's exact test was used to compare categorical variables and the results were expressed as proportions; while continuous variables were compared using the Mann-Whitney U test or student's t test and the results were described in mean \pm standard deviation (consistent with the normal distribution) or median with interquartile range (inconsistent with the normal distribution). Every statistically significant variable was analyzed by receiver operating characteristic curve (ROC) and then the best cutoff values of quantitative variables were obtained by Youden index for maximum specificity and sensitivity. Subsequently, the risk factors for diagnosing metastases were identified by binary logistic regression analysis. The AUC and nomograms of the original diagnostic model based on the risk factors for metastases were used to assess the diagnostic performance. According to the best cutoff values, we dichotomized the quantitative variables and then involved all risk factors into the multivariate analysis. If the risk factor was negative, the point was zero; if the risk factor was positive, the point was rounding of $\ln OR$. At last, a diagnostic scoring model was established according to the approximate value of $\ln OR$ of risk factors for metastases. The AUC and scoring table of the diagnostic scoring model were used to show the differential diagnosis ability. DeLong test was used for the comparison between AUCs. A $p < 0.05$ was treated as significant.

3 Results

3.1 Comparison of clinical and radiological features

The mean patient age of metastases (60.9 ± 8.4 years) was significantly older than that of LAPs (55.8 ± 11.7 years) ($p < 0.001$); 75.61% of metastases (93/123) were male, whereas 39.83% of LAPs

(47/118) ($p < 0.001$); 16.26% of metastases (20/123) showed irregular in shape whereas 5.93% of LAPs (7/118) did ($p = 0.011$). Cystic degeneration/necrosis was found in 25.2% of metastases (31/123) and 11.9% of LAPs (14/118), respectively ($p = 0.008$). The mean CT values in UP of metastases (37.49 ± 7.63 HU) was significantly higher than that of LAPs (27.38 ± 9.83 HU) ($p < 0.001$). The average values of ERA and ERV of metastases were 0.77 and 1.05, respectively, which were notably lower than those of LAPs (ERA: 1.60; ERV: 2.03) (all $p < 0.001$). LD, SD, Lesion location, CT values in AP, and CT values in VP were no significant differences between metastases and LAPs (all $p > 0.05$). More than half of LAPs (61.02%, 72/118) and metastases (60.98%, 75/123) were in venous phase at peak enhancement level. However, 22.03% (26/118) of LAPs were in arterial phase at peak enhancement level and only 3.25% (4/123) of metastases. While up to 35.77% (44/123) of metastases showed equally enhanced, which higher than that of LAPs. On the whole, the peak enhancement phase had significant difference between metastases and LAPs ($p < 0.001$). The highest proportion of metastases about clinical stage of lung cancer is stage IV (55.28%, 68/123), and up to 91.87% (36.59% + 55.28%) of metastases showed stages III and IV, which was significant higher than that of LAPs (61.02%, 35.60% + 25.42%); The whole clinical stage of lung cancer between LAPs and metastases was significant difference ($p < 0.001$). The proportion of small-cell lung cancer in metastases was 34.1% (42/123), which was significant higher than that of LAPs (only 8.5%, 10/118), ($p < 0.001$) (Supplementary Table S2).

The AUC values for quantitative and categorical variables with statistical significance were obtained by ROC analysis. Of these variables, the AUC of ERV was higher than that of age, CT values in UP, ERA, gender, shape, cystic degeneration, histology of lung cancer, peak enhancement phase, and clinical stage of lung cancer (Table 1; Figure 2). The cutoff values for age, CT values in UP, ERA, and ERV were 59.5 year, 27.5 HU, 1.04, and 1.44, respectively (Table 1).

3.2 Multivariate logistic regression analysis

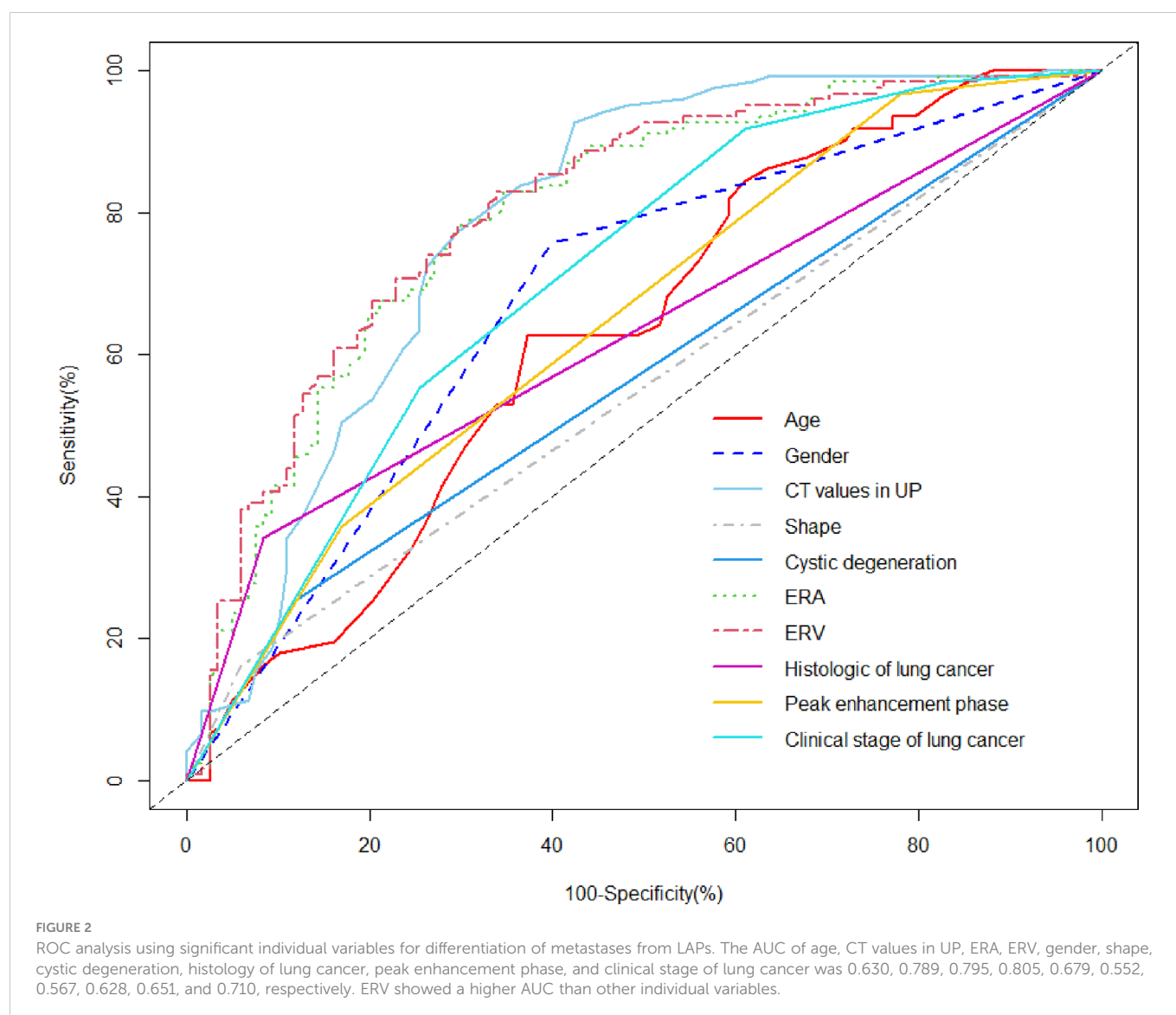
Ten variables—age, gender, CT values in UP, shape, cystic degeneration, ERA, ERV, peak enhancement phase, clinical stage of lung cancer, and histology of lung cancer based on univariate analysis—were further analyzed using multivariate logistic regression. Seven variables—age, gender, CT values in UP, cystic degeneration, ERV, peak enhancement phase, and clinical stage of lung cancer—were finally supposed to risk factors for differential diagnosis of metastasis and LAPs, and the original diagnostic model had high AUC of 0.919 (0.883–0.955) with ideal sensitivity (83.74%), specificity (88.98%), and accuracy (86.31%) (Figure 3). According to the nomogram, the probability of metastases was determined by mapping “Total Points” obtained by adding the points of each feature to the “Risk of Metastases” in the bottom of Figure 4A; the calibration curve for nomogram was showed in Figure 4B.

Table 2 showed the OR of all risk factors. Age ≥ 59.5 years (OR: 2.269; 95% CI 1.037–4.963; $p = 0.04$), male (3.511; 95% CI 1.1570–7.848; $p = 0.002$), CT values in UP ≥ 27.5 HU (OR: 6.968; 95% CI 2.580–18.824; $p < 0.001$), cystic degeneration/necrosis (OR: 3.076; 95% CI 1.110–8.521; $p = 0.031$), ERV ≤ 1.44 , (OR: 4.835; 95% CI

TABLE 1 Individual variables obtained from ROC analysis for differentiation of LAPs from metastases.

Variables	Cutoff	AUC	Sensitivity	Specificity	PPV	NPV
Age	59.5	0.630	62.6%	62.7%	63.6%	61.7%
CT values in UP	27.5	0.789	92.7%	57.6%	69.5%	88.3%
ERA	1.04	0.795	78.9%	69.5%	72.9%	75.9%
ERV	1.44	0.805	82.9%	66.1%	71.8%	78.8%
Gender	–	0.679	75.6%	60.2%	66.4%	70.3%
Shape	–	0.552	16.3%	94.1%	74.1%	51.9%
Cystic degeneration	–	0.567	25.2%	88.1%	68.9%	53.1%
Histology of lung cancer	–	0.628	34.1%	91.5%	80.8%	57.1%
Peak enhancement phase	–	0.651	96.7%	22.3%	56.4%	86.7%
Clinical stage of lung cancer	–	0.710	91.9%	39.0%	61.1%	82.1%

ROC, receiver operating characteristic curve; UP, unenhanced phase; ERA, enhancement ratio on arterial phase; ERV, enhancement ratio on venous phase; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.



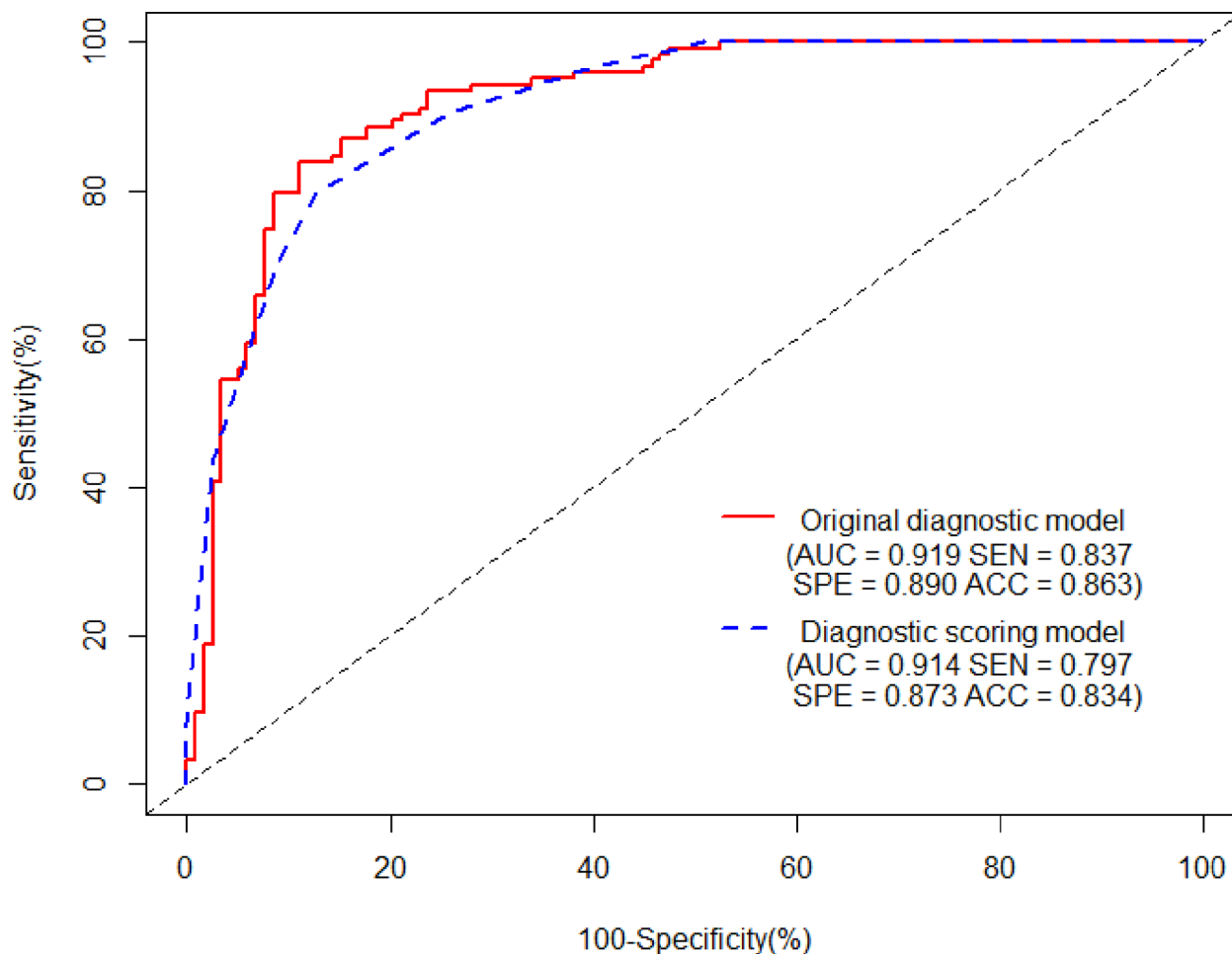


FIGURE 3

The AUCs of the original diagnostic model and the diagnostic scoring model for metastases were 0.919 (0.883–0.955) and 0.914 (0.880–0.948), respectively. There was no statistical significance of AUC between the two diagnostic model ($P = 0.644$). As for the diagnostic scoring model, a cutoff value of ≥ 9 points yielded a sensitivity of 89.43% and a specificity of 76.27% for diagnosis of metastases.

2.006–11.694; $p < 0.001$), venous phase or equally enhanced (OR: 16.907; 95% CI 4.319–66.182; $p < 0.001$ or OR: 14.036; 95% CI 3.429–57.455; $p < 0.001$), and clinical stage II or III or IV (OR: 3.550; 95% CI 0.496–25.433; $p = 0.208$ or OR: 17.535; 95% CI 2.959–103.905; $p = 0.002$ or OR: 20.241; 95% CI 3.532–116.007; $p = 0.001$) were all risk factors for the diagnosis of metastases. According to the OR of all hazard variables, a diagnostic scoring system was established and the specific score assignment was 1 point for age ≥ 59.5 years and male and cystic degeneration/necrosis and stage II, 2 points for CT values in UP ≥ 27.5 HU and ERV ≤ 1.44 , and 3 points for venous phase or equally enhanced and stage III or stage IV. The diagnostic scoring model also performed relatively high AUC of 0.914 (0.880–0.948) with a sensitivity of 79.67%, a specificity of 87.29%, and an accuracy of 83.40%. There was no statistical significance of AUC between the two diagnostic model ($p = 0.644$), which indicated that the diagnostic scoring model was simplified but did not affect the discriminative accuracy of metastases (Figure 3). Table 3 showed the diagnostic performances with the best cutoff values. The sensitivity was 79.67% and the specificity was 87.29% when the best cutoff value ≥ 10 points. In addition, LAPs was highly hinted with the possibility as high as 93.6% when the nodule's score was less than 8 points; metastases was

highly hinted with the possibility as high as 94.7% if the score was equal to or greater than 12 points (Table 3). Meanwhile, the AUC of the diagnostic scoring model was significantly higher than every individual variables for differentiate metastases from LAPs (Table 4). Examples are given in Figures 5, 6.

Discussion

Globally, lung cancer is the main cause of cancer death (25). Almost half of lung cancer patients are found to have distant metastasis at the time of diagnosis and a frequent site of metastatic spread is the adrenal gland. Meanwhile, adrenal adenomas also have a high frequency (approximately 9%) in general population. No single imaging examination method can be regarded as the gold standard based on current studies to differentiate metastases from adenomas, especially in lung cancer cases (26). Owing to demanding time-consuming and computer expertise for the analysis of high-dimensional features that cannot be recognized by the naked eye, radiomics, as advanced image analysis techniques, has not been widely used in clinical practice. Thus, the examination and follow-

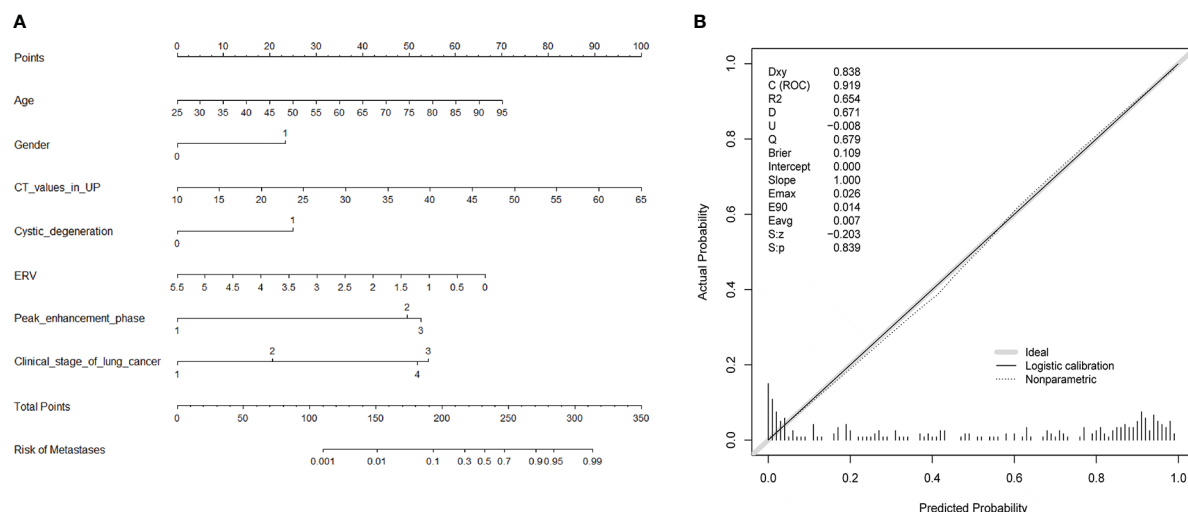


FIGURE 4
The nomogram of the original diagnostic model (A). The calibration curve for nomogram (B).

TABLE 2 Multivariate regression analysis for identifying metastases.

Variables	OR (95% CI)	P value
Age		0.040
< 59.5	1.0	
≥ 59.5	2.269 (1.031–4.963)	
Gender		0.002
Female	1.0	
Male	3.511 (1.570–7.848)	
CT values in UP		< 0.001
< 27.5	1.0	
≥ 27.5	6.968 (2.580–18.824)	
Cystic degeneration/necrosis		0.031
No	1.0	
Yes	3.076 (1.110–8.521)	
ERV		< 0.001
> 1.44	1.0	
≤ 1.44	4.835 (2.006–11.694)	
Peak enhancement phase		
Arterial phase	1.0	< 0.001
Venous phase	16.907 (4.319–66.182)	< 0.001
Equally enhanced	14.036 (3.429–57.455)	< 0.001
Clinical stage of lung cancer		
I	1.0	0.001
II	3.550 (0.496–25.433)	0.208
III	17.535 (2.959–103.905)	0.002
IV	20.241 (3.532–116.007)	0.001

OR, odds ratio; CI, confidence interval; UP, unenhanced phase; ERV, enhancement ratio on venous phase.

TABLE 3 Diagnostic performance of the diagnostic score model with different cutoffs for metastases.

Cutoff	AUC	Sensitivity(%)	Specificity(%)	PPV(%)	NPV(%)
≥2	0.5(0.5–0.5)	100(97.0–100)	0(0.0–3.1)	51.0(51.0–51.0)	
≥3	0.504(0.496–0.513)	100(97.0–100)	0.85(0.02–4.6)	51.2(50.8–51.7)	100
≥4	0.538(0.514–0.562)	100(97.0–100)	7.63(3.5–14.0)	53.0(51.7–54.3)	100
≥5	0.602(0.565–0.638)	100(97.0–100)	20.34(13.5–28.7)	56.7(54.4–58.9)	100
≥6	0.644(0.603–0.685)	100(97.0–100)	28.81(20.8–37.9)	59.4(56.6–62.2)	100
≥7	0.746(0.701–0.791)	100(97.0–100)	49.15(39.8–58.5)	67.2(63.2–71.0)	100
≥8	0.789(0.742–0.836)	95.93(90.8–98.7)	61.86(52.5–70.6)	72.4(67.5–76.8)	93.6(85.9–97.2)
≥9	0.824(0.777–0.872)	90.24(83.6–94.9)	74.58(65.7–82.1)	78.7(73.0–83.5)	88.0(80.9–92.7)
≥10	0.835(0.788–0.882)	79.67(71.5–86.4)	87.29(79.9–92.7)	86.7(80.2–91.4)	80.5(74.3–85.5)
≥11	0.803(0.755–0.851)	69.11(60.1–77.1)	91.53(85.0–95.9)	89.5(82.3–94.0)	74.0(68.5–78.8)
≥12	0.707(0.661–0.753)	43.90(35.0–53.1)	97.46(92.7–99.5)	94.7(85.3–98.2)	62.5(58.7–66.1)
≥13	0.541(0.516–0.565)	8.13(4.0–14.4)	100(96.9–100)	100	51.1(49.8–52.4)
≥14	0.5 (0.5–0.5)	0(0.0–3.0)	100(96.9–100)		49.0(49.0–49.0)

Numbers in the parentheses were 95% confidence interval; PPV, positive predictive value; NPV, negative predictive value.

up of adrenal nodules still rely mainly on the conventional imaging characteristics by visual evaluation. We evaluated the value of traditional biphasic CECT in differentiation of metastases from LAPs in this study and found that age, gender, cystic degeneration/necrosis, CT values in UP, ERV, peak enhancement phase, and clinical stage of lung cancer were risk factors for distinguishing metastases. The diagnostic scoring model established by using the above risk factors had robust ability to distinguish metastases from LAPs with AUC of 0.914 (0.880–0.948). Furthermore, compared with the original diagnostic model (or nomogram), the diagnostic scoring model (or table) had considerable differential diagnostic ability and great prospects in clinical implication because of its more simple, convenient, and accurate in the use.

We have integrated the conventional clinical and imaging variables to improve the differential diagnostic ability of biphasic CECT. Age, gender and clinical stage of lung cancer, as clinical risk variables, were all independent factors for differentiating metastases

of lung cancer from LAPs in our study. Compared with LAPs, patients with metastases were more likely to be male and older. Male and a cutoff value of ≥ 59.5 years showed a higher probability for diagnosing metastases of lung cancer, which may be related to the selection of lung cancer as the primary cancer (27). Among the adrenal metastases of lung cancer, the proportions for clinical stages I and II were only 1.63% and 6.5%, respectively, whereas 36.59% and 55.28% for stages III and IV, respectively. The results were consistent with prior study (28). The later the clinical stage of lung cancer was, the higher probability adrenal nodular was to be diagnosed as metastases.

CT values in UP, the presence of cystic degeneration or necrosis, ERV, and peak enhancement phase, as CT imaging risk variables, were all independent factors for differentiating metastases of lung cancer from LAPs in our study. Compared with LAPs, metastases showed significantly higher CT values in UP and a cutoff of ≥ 28.5 HU with an OR of 1.179 in our study, consistent with the finding of Ho et al. (29), who found that there was no statistical significance of CT values in UP between benign and

TABLE 4 Comparison of performance of the diagnostic scoring model and individual variables for differentiate metastases from LAPs.

The scoring model vs. individual variables	AUC	Z statistic	p
Scoring model vs. age	0.914 vs. 0.630	7.522	< 0.001
Scoring model vs. CT values in UP	0.914 vs. 0.789	4.345	< 0.001
Scoring model vs. ERA	0.914 vs. 0.795	4.626	< 0.001
Scoring model vs. ERV	0.914 vs. 0.805	4.168	< 0.001
Scoring model vs. gender	0.914 vs. 0.679	7.488	< 0.001
Scoring model vs. shape	0.914 vs. 0.552	13.712	< 0.001
Scoring model vs. cystic degeneration/necrosis	0.914 vs. 0.567	11.882	< 0.001
Scoring model vs. histology of lung cancer	0.914 vs. 0.628	10.164	< 0.001
Scoring model vs. peak enhancement phase	0.914 vs. 0.651	8.777	< 0.001
Scoring model vs. clinical stage of lung cancer	0.914 vs. 0.710	7.074	< 0.001

AUC, area under the curve.

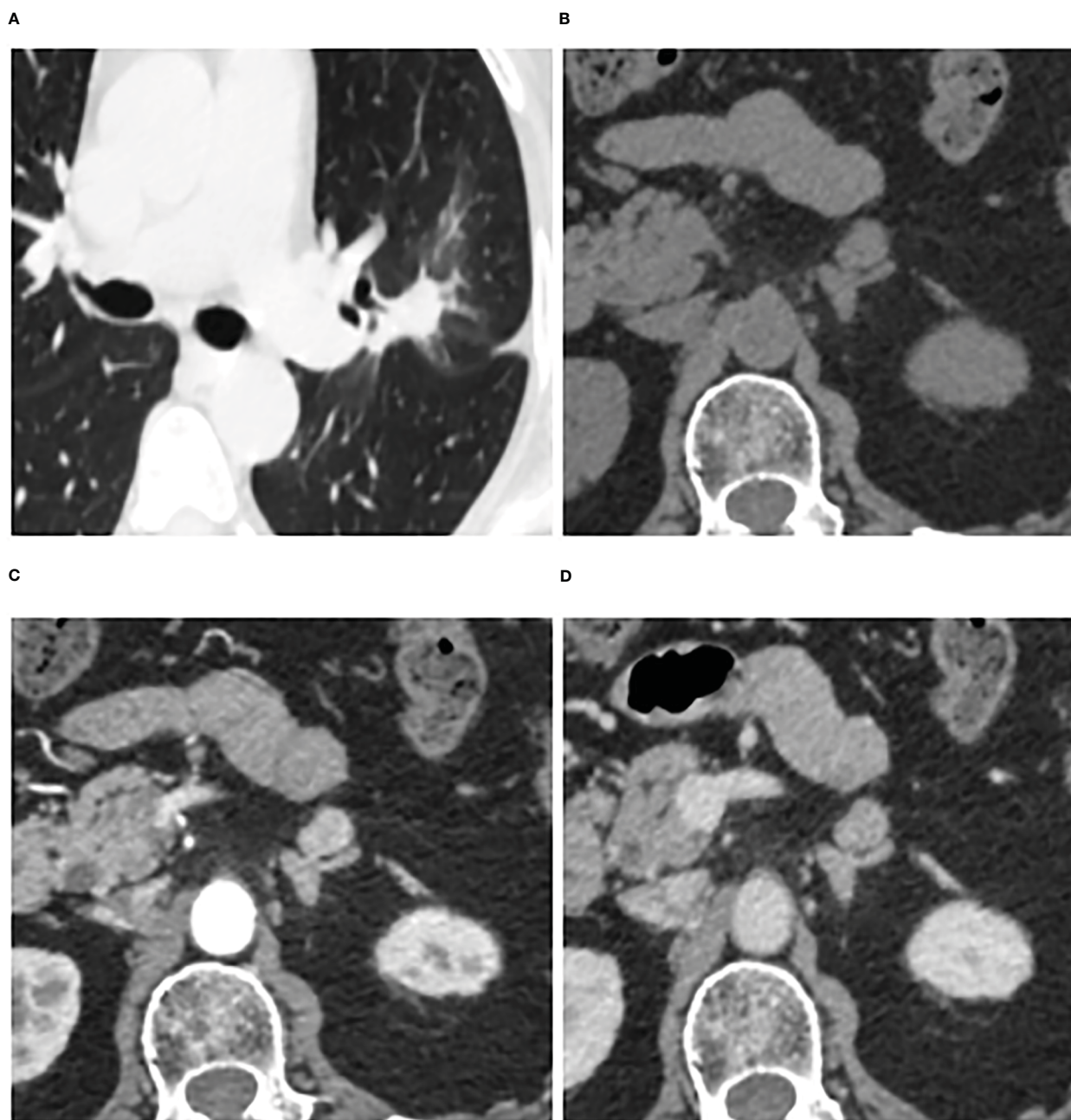


FIGURE 5

Adrenal adenoma in a 72-year-old woman. Pulmonary window of primary nonsmall-cell lung cancer (stage IV) (A) and unenhanced phase (B), arterial phase (C), and venous phase (D) of adrenal nodular showed a 13 mm × 11 mm tumor in left adrenal. Attenuation values on unenhanced, arterial, and venous phase were 39, 127, and 101 HU, respectively. The enhancement ratio on venous phase was 1.59. No cystic degeneration was seen within the nodular. Peak enhancement phase was arterial phase. The nodular got a score of 6 points, indicating diagnosis of adenoma.

malignant adrenal lesions. Additionally, cystic degeneration or necrosis, reflecting the biological behavior and structural features of masses to a certain extent, was another key variable for differentiation between metastases and LAPs. The probability of cystic degeneration or necrosis in metastases was higher than LAPs, in line with previous studies (29, 30). Song et al. showed cystic degeneration or necrosis had a sensitivity of 20–55% and a specificity of 86–93% in distinguishing adrenal metastatic nodules from benign lesions (mainly adenomas) (30). The wash-in characteristics of adrenal lesions could be reflected by ERA and ERV. The ERA and ERV of metastases were notably lower than that of LAPs, and ERV was a hazard characteristic for differentiating metastasis proved by

multivariate analysis (OR: 4.835; 95% CI: 2.006–11.694) in our study. The results were consistent with Lee et al. (21) who revealed that ERV of lung cancer metastases was notably lower than that of hyperattenuating benign lesions. However, Foti et al. (20) reported that the ERV of adenomas was not significantly different from metastases. The contradiction between studies may be due to differences in study cohort. First, previous studies had relatively small sample size. Second, all lesions were small (LD < 3 cm) and adenomas were all LAPs and metastases were all of lung cancer in our study. 61.02 and 22.03% of LAPs performed peak enhancement level in venous phase and arterial phase in our study, respectively, whereas 60.98% of metastases performed peak enhancement level in venous phase and only

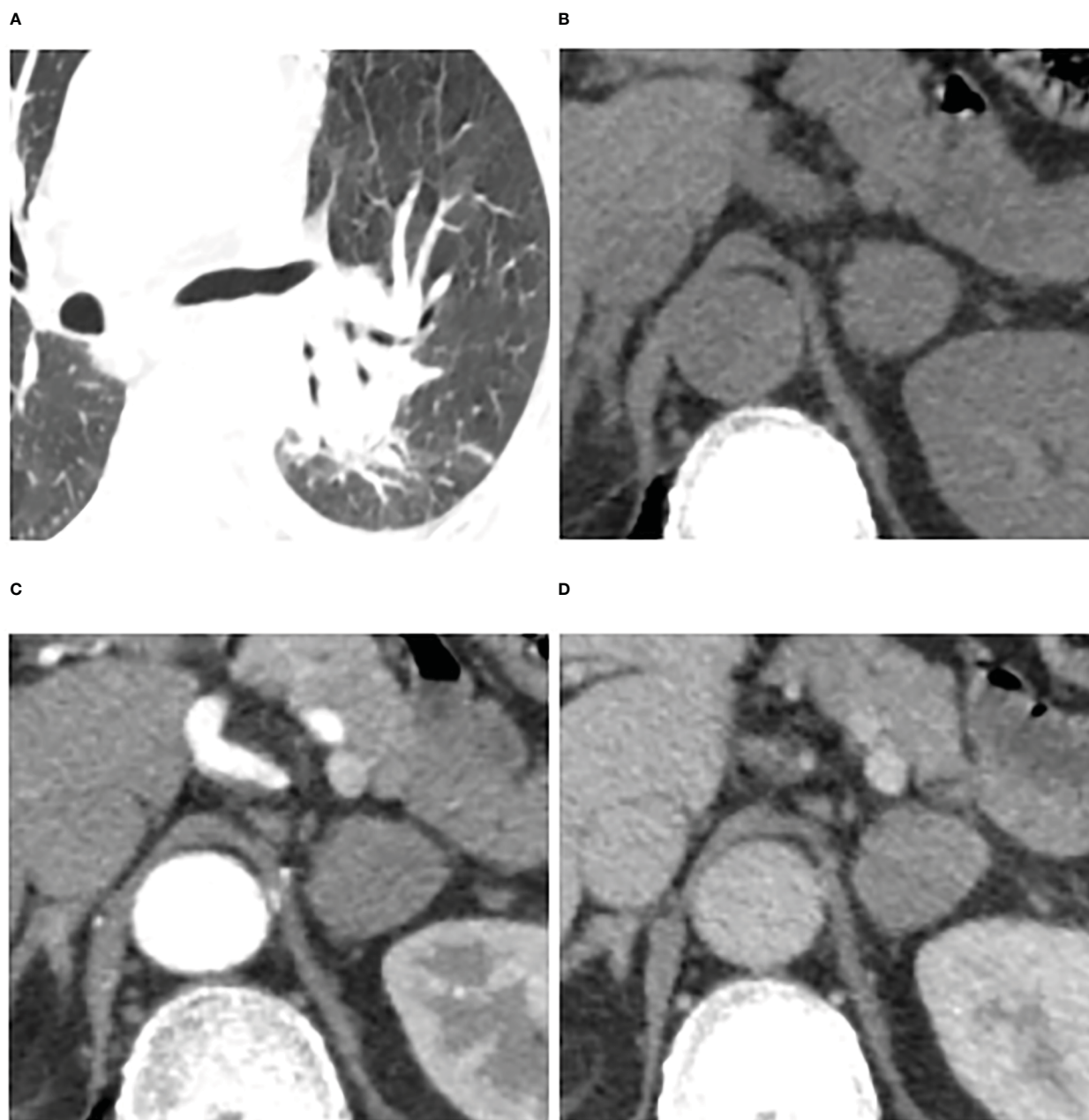


FIGURE 6

Adrenal metastases in a 66-year-old man. Pulmonary window of primary small-cell lung cancer (stage IV) (A) and unenhanced phase (B), arterial phase (C), and venous phase (D) of adrenal nodular showed a 25 mm x 23 mm tumor in left adrenal. Attenuation values on unenhanced, arterial, and venous phase were 36, 65, and 64 HU, respectively. The enhancement ratio on venous phase was 0.78. No cystic degeneration was seen within the nodular. Peak enhancement phase was equally enhanced. The nodular got a score of 12 points, indicating diagnosis of metastases.

3.25% in arterial phase and up to 35.77% equally enhanced, meaning that LAPs show relatively faster wash-in and earlier wash-out enhancement pattern than metastases, in line with previous studies (20, 31). Foti et al. (20) found that adenomas had significantly faster wash-in enhancement pattern than metastases. An et al. (31) also found up to 68.2% of LAPs had peak enhancement level in venous phase. Further studies are necessary for the difference of peak enhancement phase between LAPs and metastases of lung cancer on biphasic CECT.

Previous studies only paid close attention to individual CT radiological characteristics for differentiating metastases from adenomas. We discovered that the diagnostic accuracy was

remarkably improved by combining these clinical and imaging characteristics, although these characteristics have low diagnostic specificity when used alone. The advantage of this research was that we identified independent risk factors based on large sample by multivariate analysis and established a diagnostic scoring model based on the OR by integrating above risk features. The diagnostic scoring mode performed well in differential diagnosis of metastases from LAPs with an AUC of 0.914 (0.880–0.948), which was significantly higher than that of each individual variable (p all < 0.05). In addition, although the performances of the original diagnostic model (nomogram) and the diagnostic scoring model

(table) were similar ($p > 0.05$), we needed to obtain the points of each risk factor by visual comparison when using nomogram to predict the risk of metastases, which was prone to errors, resulting in inaccurate total points and inaccurate final prediction probability. We could directly get the points of each risk factor in the diagnostic scoring table without visual comparison. Therefore, the scoring table is more simple, convenient, and accurate in the use than nomogram, which provides an additional choice for peers.

This study had several limitations. First, our study may have some selection bias because of the retrospective nature. Second, two CT scanners and chest or abdominal CECT were used owing to the retrospective nature of our study. However, it can be considered as the strength of this research, because it conforms to the fact in work practice and provides certain potential generalizability. Third, because of lacking the delay phase, we did not estimated washout features.

In a word, biphasic CECT had well-diagnostic ability in differentiating metastases from LAPs. Age, gender, the presence of cystic degeneration, CT values in UP, ERV, peak enhancement phase, and clinical stage of lung cancer were independent risk factors for distinguishing metastases. A diagnostic scoring model integrating these risk factors with an AUC of 0.914 is easy to popularize due to simplicity and convenience.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Tangshan People's Hospital Institutional Ethics Committee. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

References

1. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. *Int J Cancer* (2021) 149:778–9. doi: 10.1002/ijc.33588
2. Buero A, Nardi W, Chimondeguy DJ, Pankl LG, Lyons GA, Arboit DG, et al. Outcomes of surgical treatment for isolated adrenal metastasis from non-small cell lung cancer. *Ecancermedicalscience* (2021) 15:1322–9. doi: 10.3332/ecancer.2021.1322
3. Almeida MQ, Bezerra-Neto JE, Mendonça BB, Latronico AC, Fragoso MCBV. Primary malignant tumors of the adrenal glands. *Clinics (Sao Paulo)* (2018) 73:e756s. doi: 10.6061/clinics/2018/e756s
4. Klikovits T, Lohinai Z, Fábán K, Gyulai M, Szilasi M, Varga J, et al. New insights into the impact of primary lung adenocarcinoma location on metastatic sites and sequence: A multicenter cohort study. *Lung Cancer* (2018) 126:139–48. doi: 10.1016/j.lungcan.2018.11.004
5. Blažeković I, Jukić T, Granić R, Punda M, Franceschi M. An unusual case of papillary thyroid carcinoma iodine-131 avid metastasis to the adrenal gland. *Acta Clin Croat* (2018) 57:372–6. doi: 10.20471/acc.2018.57.02.20
6. Badawy AA, Khedr G, Omar A, Bae S, Arafat W, Grant S. Site of metastases as prognostic factors in unselected population of stage IV non-small cell lung cancer. *Asian Pac J Cancer Prev* (2018) 19:1907–10. doi: 10.22034/APJCP.2018.19.7.1907
7. Young WF. Clinical practice. the incidentally discovered adrenal mass. *N Engl J Med* (2007) 356:601–10. doi: 10.1056/NEJMcp065470
8. Ilias I, Sahdev A, Reznick RH, Grossman AB, Pacak K. The optimal imaging of adrenal tumours: A comparison of different methods. *Endocr Relat Cancer* (2007) 14:587–99. doi: 10.1677/ERC-07-0045
9. Zeiger MA, Siegelman SS, Hamrahian AH. Medical and surgical evaluation and treatment of adrenal incidentalomas. *J Clin Endocrinol Metab* (2011) 96:2004–15. doi: 10.1210/jc.2011-0085
10. Mayo-Smith WW, Song JH, Boland GL, Francis IR, Israel GM, Mazzaglia PJ, et al. Management of incidental adrenal masses: A white paper of the ACR incidental findings committee. *J Am Coll Radiol* (2017) 14:1038–44. doi: 10.1016/j.jacr.2017.05.001
11. Bednarczuk T, Bolanowski M, Sworczak K, Górnicka B, Cieszanowski A, Otto M, et al. Adrenal incidentaloma in adults - management recommendations by the polish society of endocrinology. *Endokrynol Pol* (2016) 67:234–258. doi: 10.5603/EP.a2016.0039
12. Pandharipande PV, Herts BR, Gore RM, Mayo-Smith WW, Harvey HB, Megibow AJ, et al. Rethinking normal: Benefits and risks of not reporting harmless incidental findings. *J Am Coll Radiol* (2016) 13:764–7. doi: 10.1016/j.jacr.2016.03.017
13. Fujiyoshi F, Nakajo M, Fukukura Y, Tsuchimochi S. Characterization of adrenal tumors by chemical shift fast low-angle shot MR imaging: comparison of four methods of quantitative evaluation. *AJR* (2003) 180:1649–57. doi: 10.2214/ajr.180.6.1801649
14. Guerin C, Pattou F, Brunaud L, Lifante JC, Mirallié E, Haissaguerre M, et al. Performance of 18F-FDG PET/CT in the characterization of adrenal masses in noncancer

Author contributions

WX, and LC together designed the study. LC collected the patient images and performed the statistical analysis and wrote the manuscript. WX and LZ critically reviewed the manuscript, LZ also provided statistical guidance. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to thank Wengui Xu for the study design and editing the draft of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1091102/full#supplementary-material>

patients: a prospective study. *J Clin Endocrinol Metab* (2017) 102:2465–72. doi: 10.1210/jc.2017-00254

15. Caoili EM, Korobkin M, Francis IR, Cohan RH, Platt JF, Dunnick NR, et al. Adrenal masses: Characterization with combined unenhanced and delayed enhanced CT. *Radiology* (2002) 222:629–33. doi: 10.1148/radiol.2223010766

16. Choi YA, Kim CK, Park BK, Kim B. Evaluation of adrenal metastases from renal cell carcinoma and hepatocellular carcinoma: Use of delayed contrast-enhanced CT. *Radiology* (2013) 266:514–20. doi: 10.1148/radiol.12120110

17. Koo HJ, Choi HJ, Kim HJ, Kim SO, Cho KS. The value of 15-minute delayed contrast-enhanced CT to differentiate hyper attenuating adrenal masses compared with chemical shift MR imaging. *Eur Radiol* (2014) 24:1410–20. doi: 10.1007/s00330-013-3084-7

18. Akkus G, Güney IB, Ok F, Evran M, Izol V, Erdoğan S, et al. Diagnostic efficacy of 18F-FDG PET/CT in patients with adrenal incidentaloma. *Endocr Connect* (2019) 8:838–45. doi: 10.1530/EC-19-0204

19. Kassirer JP. Our stubborn quest for diagnostic certainty. *N Engl J Med* (1989) 320:1489–91. doi: 10.1056/NEJM198911023211820

20. Foti G, Faccioli N, Manfredi R, Mantovani W, Mucelli RP. Evaluation of relative wash-in ratio of adrenal lesions at early biphasic CT. *AJR Am J Roentgenol* (2010) 194:1484–91. doi: 10.2214/AJR.09.3636

21. Lee HY, Oh YL, Park SY. Hyperattenuating adrenal lesions in lung cancer: Biphasic CT with unenhanced and 1-min enhanced images reliably predicts benign lesions. *Eur Radiol* (2021) 31:5948–58. doi: 10.1007/s00330-020-07648-1

22. Nishino M, Jagannathan JP, Ramaiya NH, Abbeele ADV. Revised RECIST guideline version 1.1: What oncologists want to know and what radiologists need to know. *AJR Am J Roentgenol* (2010) 195:281–9. doi: 10.2214/AJR.09.4110

23. Northcutt BG, Raman SP, Long C, Oshmyansky AR, Siegelman SS, Fishman EK, et al. MDCT of adrenal masses: Can dual-phase enhancement patterns be used to

differentiate adenoma and pheochromocytoma? *AJR Am J Roentgenol* (2013) 201:834–9. doi: 10.2214/AJR.12.9753

24. Goroshi M, Jadhav SS, Sarathi V, Lila AR, Patil VA, Shah R, et al. Radiological differentiation of pheochromocytoma from other malignant adrenal masses: Importance of wash-in characteristics on multiphase CECT. *Endocr Connect* (2019) 8:898–905. doi: 10.1530/EC-19-0198

25. Simsek FS, Arslan M, Dag Y. An exceptional group of non-small cell lung cancer difficult to diagnose: Evaluation of lipid-poor adrenal lesions. *Bosn J Basic Med Sci* (2019) 19:195–200. doi: 10.17305/bjbm.2019.3837

26. Cingam SR, Mukkamalla SKR, Karanchi H. *Adrenal metastasis*. Treasure Island (FL: StatPearls Publishing (2022).

27. Groot PD, Munden RF. Lung cancer epidemiology, risk factors, and prevention. *Radiol Clin North Am* (2012) 50:863–76. doi: 10.1016/j.rcl.2012.06.006

28. Wood SL, Pernemalm M, Crosbie PA, Whetton AD. The role of the tumor-microenvironment in lung cancer-metastasis and its relationship to potential therapeutic targets. *Cancer Treat Rev* (2014) 40:558–66. doi: 10.1016/j.ctrv.2013.10.001

29. Ho LM, Samei E, Mazurowski MA, Zheng Y, Allen BC, Nelson RC, et al. Can texture analysis be used to distinguish benign from malignant adrenal nodules on unenhanced CT, contrast-enhanced CT, or InPhase and opposed-phase MRI? *AJR Am J Roentgenol* (2019) 212:554–61. doi: 10.2214/AJR.18.20097

30. Song JH, Grand DJ, Beland MD, Chang KJ, Machan JT, Mayo-Smith WW. Morphologic features of 211 adrenal masses at initial contrast-enhanced CT: Can we differentiate benign from malignant lesions using imaging features alone? *AJR* (2013) 201:1248–53. doi: 10.2214/AJR.12.10302

31. An YY, Yang GZ, Lin B, Zhang N, Hou HT, Zhu FM, et al. Differentiation of lipid-poor adenoma from pheochromocytoma on biphasic contrast-enhanced CT. *Abdom Radiol (NY)* (2021) 46:4353–61. doi: 10.1007/s00261-021-03121-9



OPEN ACCESS

EDITED BY
Alla Reznik,
Lakehead University, Canada

REVIEWED BY
Manuel Conson,
University of Naples Federico II, Italy
Joost Nederend,
Catharina Hospital, Netherlands

*CORRESPONDENCE
Fenhua Zhao
✉ zhfenhua@163.com
Jihong Sun
✉ sunjihong@zju.edu.cn
Long Zhou
✉ zhoulong21@zju.edu.cn

†These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 24 October 2022
ACCEPTED 06 February 2023
PUBLISHED 20 February 2023

CITATION
Fu C, Shao T, Hou M, Qu J, Li P, Yang Z,
Shan K, Wu M, Li W, Wang X, Zhang J,
Luo F, Zhou L, Sun J and Zhao F (2023)
Preoperative prediction of tumor deposits
in rectal cancer with clinical-magnetic
resonance deep learning-based
radiomic models.
Front. Oncol. 13:1078863.
doi: 10.3389/fonc.2023.1078863

COPYRIGHT
© 2023 Fu, Shao, Hou, Qu, Li, Yang, Shan,
Wu, Li, Wang, Zhang, Luo, Zhou, Sun and
Zhao. This is an open-access article
distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Preoperative prediction of tumor deposits in rectal cancer with clinical-magnetic resonance deep learning-based radiomic models

Chunlong Fu^{1†}, Tingting Shao^{2†}, Min Hou^{2†}, Jiali Qu², Ping Li^{2,3},
Zebin Yang¹, Kangfei Shan¹, Meikang Wu¹, Weida Li¹,
Xuan Wang², Jingfeng Zhang⁴, Fanghong Luo², Long Zhou^{2*},
Jihong Sun^{2,4,5*} and Fenhua Zhao^{1*}

¹Department of Radiology, Affiliated Dongyang Hospital of Wenzhou Medical University, Dongyang, China, ²Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, ³Department of Radiology, Jiaying Hospital of Traditional Chinese Medicine, Jiaying, China, ⁴Key Laboratory of Diagnosis and Treatment of Digestive System Tumors of Zhejiang Province, Ningbo, China, ⁵Cancer Center, Zhejiang University, Hangzhou, China

Background: This study aimed to establish an effective model for preoperative prediction of tumor deposits (TDs) in patients with rectal cancer (RC).

Methods: In 500 patients, radiomic features were extracted from magnetic resonance imaging (MRI) using modalities such as high-resolution T2-weighted (HRT2) imaging and diffusion-weighted imaging (DWI). Machine learning (ML)-based and deep learning (DL)-based radiomic models were developed and integrated with clinical characteristics for TD prediction. The performance of the models was assessed using the area under the curve (AUC) over five-fold cross-validation.

Results: A total of 564 radiomic features that quantified the intensity, shape, orientation, and texture of the tumor were extracted for each patient. The HRT2-ML, DWI-ML, Merged-ML, HRT2-DL, DWI-DL, and Merged-DL models demonstrated AUCs of 0.62 ± 0.02 , 0.64 ± 0.08 , 0.69 ± 0.04 , 0.57 ± 0.06 , 0.68 ± 0.03 , and 0.59 ± 0.04 , respectively. The clinical-ML, clinical-HRT2-ML, clinical-DWI-ML, clinical-Merged-ML, clinical-DL, clinical-HRT2-DL, clinical-DWI-DL, and clinical-Merged-DL models demonstrated AUCs of 0.81 ± 0.06 , 0.79 ± 0.02 , 0.81 ± 0.02 , 0.83 ± 0.01 , 0.81 ± 0.04 , 0.83 ± 0.04 , 0.90 ± 0.04 , and 0.83 ± 0.05 , respectively. The clinical-DWI-DL model achieved the best predictive performance (accuracy 0.84 ± 0.05 , sensitivity 0.94 ± 0.13 , specificity 0.79 ± 0.04).

Conclusions: A comprehensive model combining MRI radiomic features and clinical characteristics achieved promising performance in TD prediction for RC patients. This approach has the potential to assist clinicians in preoperative stage evaluation and personalized treatment of RC patients.

KEYWORDS

deep learning, rectal cancer, tumor deposit, magnetic resonance imaging, diffusion-weighted imaging

1 Introduction

Colorectal cancer (CRC) is the third most common malignancy and second leading cause of death worldwide. In particular, rectal cancer (RC) accounts for approximately one-third of CRC (1). A tumor deposit (TD) is defined as a discontinuous cancerous nodule located in the mesenteric fascia without obvious nodal or vascular features. The median incidence of TDs in CRC patients is approximately 21.3% (2). Positive TDs can elevate clinical stages of RC patients. RC patients with positive TDs are classified as N1c and treated as clinical stage III, in the absence of nodal metastases. The efficacy of adjuvant chemotherapy in stage III colon cancer had been widely recognized in previous studies. For the TD patients, with the receipt of chemotherapy had decreased risk of cancer-specific mortality compared with those not (3). However, only 52% of TD-positive and lymph node (LN)-negative patients receive preoperative adjuvant chemotherapy (4). Therefore, early identification of TDs is important and valuable for stage evaluation and treatment planning.

Magnetic resonance imaging (MRI) is considered the most reliable imaging modality for the initial pretreatment evaluation of patients with RC, including the assessment of TN staging, circumferential resection margin (CRM), and extramural vascular invasion (EMVI). Moreover, it can assist in the clinical preoperative management of RC patients, determination of surgical scope, and assessment of treatment response to neoadjuvant therapy (5, 6). However, MRI is limited in determining tumor spread in the mesorectum. Gröne et al. used a small diameter of 5 mm as the statistical threshold to determine N staging. With this value, the sensitivity, specificity, and accuracy of MRI staging were 72%, 45.7%, and 56.7%, respectively (7). Langman et al. showed that mesenteric nodules <3 mm had a 28% probability of malignancy (8). These studies focused on the presence of tumor spread in mesorectal nodules, either LN metastasis or TDs. However, the evaluation of malignant LNs alone is insufficient to reflect the actual spread of RC in the mesentery. According to the current European Society for Medical Oncology (ESMO) preoperative risk assessment criteria, patients with TDs are classified into a high-risk group with a worse prognosis (9). A previous study confirmed that the presence of TDs is an independent risk factor for the prognosis of patients with RC (10). An analysis of two prognostic studies in N0 and N1c stages showed a significant difference in the five-year survival rates (N0, 91.5%; N1c, 37%) (11). Therefore, preoperative determination of tumor deposition status in patients with RC is essential for optimal treatment.

Currently, the presence of TDs is determined by pathological analysis after radical tumor resection. However, this method is invasive and can be performed only postoperatively. According to a previous study, MRI can help preoperatively and identify TDs and LN metastases, as these lesions appear to have distinguishable imaging characteristics on MRI (12). However, TDs are usually less than 5 mm in diameter, and identifying such small nodules and accurately assess the characteristics of the nodule can be challenging for radiologists who are already overburdened in reading MRI in daily practice.

By extracting vast amounts of quantitative features from imaging and providing non-visual information that indicates the biological behavior of tumors, radiomics has gained popularity for the non-invasive prediction of clinical or prognostic features of tumors, such as T staging of RCs, LN status, vascular and nerve invasion, distant metastasis, and pathological complete response to neoadjuvant chemotherapy (13–20). Meanwhile, artificial intelligence including deep neural networks has demonstrated high performance in the analysis of medical images (21–23), providing cancer risk assessment, recurrence, and survival predictions with higher accuracy than human experts. Recently, several radiomic models have been developed based on ultrasound (US), computed tomography (CT), and MRI to preoperatively predict TDs in patients with RC (10, 24, 25). However, the sample sizes in these studies were relatively small. Furthermore, the MRI study carried out by Yang et al. (25) only extracted the radiomic features from high-resolution T2 weighted (HRT2) MRI, whereas functional MRI, such as diffusion-weighted imaging (DWI), carries more information on the heterogeneity of tumors. Currently, there is a lack of functional MRI-based deep-learning (DL) radiomics research in this field.

This study aimed to develop an MR-based DL radiomic model for preoperative TD prediction in a larger cohort with higher prediction accuracy. This model extracts radiomic features from both HRT2 and DWI images and integrates clinical factors into TD prediction.

2 Methods

2.1 Patient characteristics

The records of 784 consecutive RC patients who underwent preoperative MRI and radical surgery between 2013 and 2020 at Sir Run Run Shaw Hospital affiliated with Zhejiang University School

of Medicine were reviewed retrospectively. The local institutional review board approved this study and provided a waiver of consent. The inclusion criteria were: (a) pathologically confirmed primary RC; (b) no neoadjuvant chemotherapy or radiotherapy before surgery; and (c) tumor visible in at least three sequential slices of HRT2 MRI. The exclusion criteria were as follows: (a) inadequate MRI quality due to intractable artifacts, including HRT2 ($n = 113$) and DWI ($n = 78$); (b) tumors not visible in HRT2 images ($n = 14$); (c) carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9) levels not obtained ($n = 44$); (d) lack of tissue differentiation grading in pathology reports ($n = 5$); and (e) co-occurrence of other digestive system malignancies ($n = 30$). Ultimately, 500 patients were enrolled in this study (Figure 1).

2.2 Clinical characteristics and pathological criteria

Clinical characteristics, including sex, age, body mass index (BMI), and carbohydrate antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA) levels, were collected from electronic medical records. Histological grading, pathological tumor node metastasis (pTNM) staging, LNs, TDs, vascular and nerve invasion, and other clinical data were obtained from pathological reports. The eighth edition of the AJCC staging system was used as a reference for the pTNM staging. TDs are defined as discrete tumor foci in the pericolic or perirectal fat, without histological evidence of residual lymph node or identifiable vascular or neural structures. The distance between the tumor and anus, TN staging, CRM, and EMVI based on the MRI were obtained from standardized reports of the picture archiving and communication system (PACS). The criteria for determining LN positivity on MRI were based on the latest recommendations of the 2016 European Society of Gastrointestinal and Abdominal Radiology consensus meeting (26). MRI tumor length was defined as the T2 sagittal tumor length. The distance from the anus was

defined as the distance from the most inferior boundary of the tumor to the subcutaneous edge of the anus. The criterion for EMVI positivity was tumor invasion of the extramural vessels, with or without vessel dilatation (27). CRM positivity was defined as a tumor location within 1 mm of the mesorectal fascia, including suspicious LNs, TDs, tumor expansion, and EMVI (28). Unclear or missing information in the MRI reports were labeled and finally confirmed by a radiologist with nine years of working experience.

2.3 MRI scanning

MRI acquisitions were performed using the following 3.0-T MRI scanners: Signa HDxt (GE Healthcare, Chicago, IL), Discovery MR750w (GE Healthcare), and MAGNETOM Skyra (Siemens Healthineers, Erlangen, Germany). The MRI protocol consisted of one axial HRT2 MRI sequence and one DWI acquisition obtained using b-values of 0 and 1,000 (or 800) s/mm^2 . No intravenous contrast agents were administered. Details of the MRI acquisition parameters are listed in Table 1.

2.4 Tumor segmentation and processing

Before image segmentation, patient-sensitive information was anonymized. The primary tumor region (3D volume) was semi-manually segmented on axial HRT2 and DWI images by a junior radiologist (with more than three years of experience in radiology) using an open-source software tool (ITK-SNAP 3.8; www.itksnap.org) (29). Automatic tumor segmentation using a CE-net-based DL segmentation model (30) was performed on the axial HRT2 images to assist radiologists. All segmentation masks were reviewed by a senior radiologist (with more than five years of experience in radiology) and finally confirmed by another senior radiologist (with more than 10 years of experience in radiology). Disagreements were resolved through discussion.

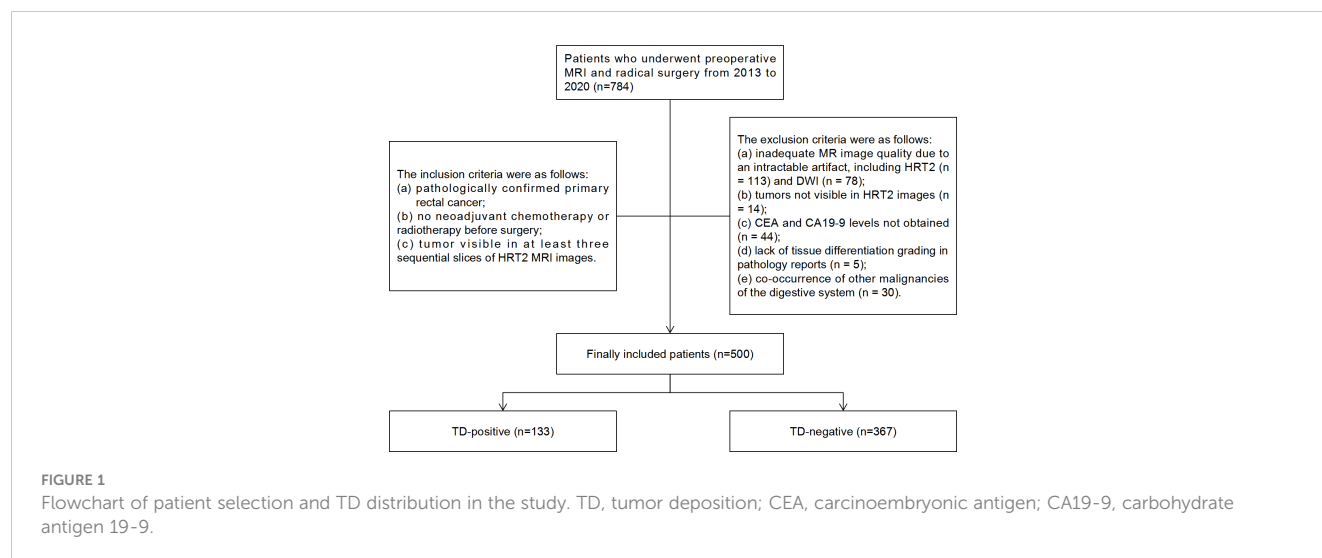


TABLE 1 Image acquisition parameters.

Machine type	GE-Signa HDxt		GE-Discovery MR750W		SIEMENS-Skya	
Modality	HRT2	DWI	HRT2	DWI	HRT2	DWI
Repetition time (ms)	3300	5900	3300	8000	5800	4400
Echo time (ms)	130	66	120	66	99	61
Slice thickness (mm)	3	5	3	5	3	5
Slice gap (mm)	0.3	1	0.3	1	0.3	1
Matrix	512×512	256×256	512×512	256×256	320×410	128×160
Echo train length	20	1	20	1	18	1
FOV (mm×mm)	160×160	250×250	160×160	380×380	160×160	300×300
b-values (s/mm2)		800		1000		800

2.5 Extraction of features

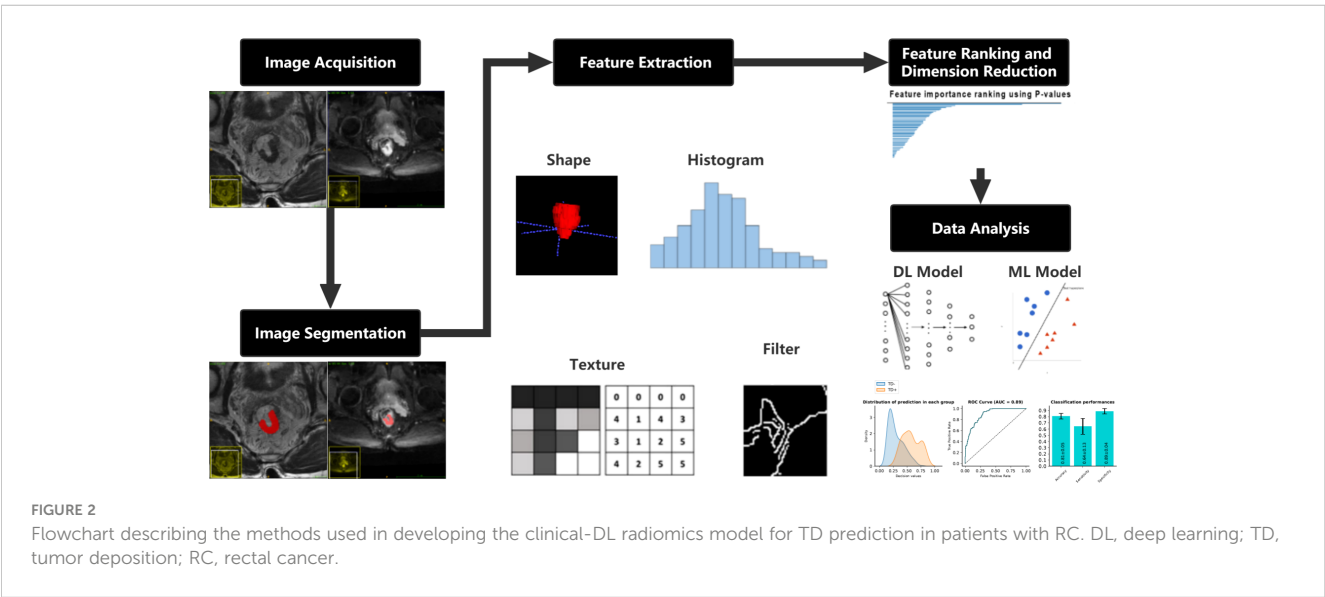
International Biomarker Standardization Initiative (IBSI)-compliant radiomic features were extracted separately for the HRT2 images and DWI images using PyRadiomics, an open-source Python package (version 2.1.2, <https://pyradiomics.readthedocs.io>) (31). Before feature extraction, z-score normalization of the MRI signal intensities for both the HRT2 and DWI images was performed using PyRadiomics. Consequently, 564 features were obtained for each HRT2 and DWI image, including 13 first-order statistics, 35 shape features, 9 orientation features, and 507 texture features, such as the gray-level co-occurrence matrix, gray-level size zone matrix, gray-level run-length matrix, gray-level dependence matrix, neighborhood gray-tone difference matrix, Gabor filter, Laplacian of Gaussian filter, local binary patterns, and local phase and vascularity filters. A variance test was performed on the extracted features to remove features with a low variance (<0.01). A t-test was used to estimate the radiomic features that were significantly correlated with TDs. Features with $p < 0.05$ were considered significant features for model development. A detailed flowchart of this process is shown in Figure 2.

2.6 Pre-processing of features

Each radiomic feature was standardized using z-score normalization to improve the robustness of the model. Missing information on the clinical characteristics was replaced with the mean value of the corresponding feature. The number of positive and negative samples was balanced using an up-sampling method within the open-source Python package Imbalanced-learn (version 0.9.0) (32).

2.7 Development of radiomic models

Radiomic models based on common machine learning (ML) techniques and DL methods were developed and compared to predict TDs. Three ML models and three DL models were constructed using features from HRT2 images (HRT2-ML and HRT2-DL models), DWI images (DWI-ML and DWI-DL models), and joint HRT2-DWI images (Merged-ML and Merged-DL models).



Integrated models combining radiomic information and clinical characteristics were developed to further improve the predictive performance. Three ML models and three DL models were constructed using clinical characteristics and features from HRT2 images (Clinical-HRT2-ML and Clinical-HRT2-DL models), DWI images (Clinical-DWI-ML and Clinical-DWI-DL models), and joint HRT2-DWI images (Clinical-Merged-ML and Clinical-Merged-DL models). For comparison, a clinical model that analyzed only clinical characteristics was also developed.

ML models used the least absolute shrinkage and selection operator (LASSO) technique (33) to select the optimized subset of features from 221 preprocessed features, followed by a support vector machine (SVM) to construct a prediction model.

DL models used a four-layer multi-layer perceptron (MLP) model, in which 221 preprocessed features were directly input. The feature numbers at each layer were 256, 128, 64, and 2, respectively, and the softmax activation function was used for the final output. The network model was actualized using the open-source deep learning framework PyTorch (34), where the batch size was set to 16, the learning rate was 0.001, and the Adam algorithm was used as the optimizer.

The predictive performance of each model was evaluated using five-fold repeated cross-validation. Each cross-validated split of the data was used to perform feature selection techniques to avoid bias in the estimation of the predictive performance.

2.8 Statistical analyses

Statistical analysis was performed with SPSS software (version 26.0; IBM, Armonk, NY) and R (version 3.5.1; R Foundation, Vienna, Austria). Differences in categorical characteristics between RC patients with and without TDs were compared using Pearson's chi-squared test and Fisher's exact test. Continuous variables are expressed as means \pm standard deviations. Differences in continuous characteristics between the two groups were compared using the Mann-Whitney U test. For all statistical analyses, $P < 0.05$ (two-sided test) was considered statistically significant. The predictive performance of the models was evaluated using the area under the receiver operating characteristic (ROC) curve (AUC), over five-fold cross-validation.

3 Results

3.1 Clinical characteristics of patients

The final 500 patients included 315 males and 185 females (mean age, 64.59 ± 10.7 years). According to the pathologically confirmed TD results, the patients were divided into TD+ ($n = 133$) and TD- ($n = 367$) subgroups. The baseline characteristics of the patients are summarized in Table 2.

3.2 Performance of the radiomic models

In total, six radiomics models were developed, and the Merge-ML and DWI-DL models both demonstrated comparable

performance, with AUCs of 0.69 ± 0.04 and 0.68 ± 0.03 , respectively ($P < 0.05$). The other models' AUCs were lower than those of the aforementioned two models. The AUCs of the HRT2-ML, HRT2-DL, DWI-ML, and Merged-DL models were 0.62 ± 0.02 , 0.57 ± 0.06 , 0.64 ± 0.08 and 0.59 ± 0.04 , respectively.

3.3 Performance of the integrated models combining radiomic information and clinical characteristics

TN staging, tumor length (measured in the sagittal view), tumor index CEA, CRM, and EMVI as assessed in the MRI report were significantly different between the TD+ and TD- groups. These clinical markers were used to establish a clinical model and were introduced into the integrated models. The performance of all models are listed in Table 3. Both the Clinical-ML and Clinical-DL models performed similarly, with AUC values of 0.81 ± 0.04 and 0.81 ± 0.06 , respectively. Among the integrated models, the Clinical-DWI-DL model achieved the highest performance, with a diagnostic accuracy of 0.84 ± 0.05 , an AUC score of 0.90 ± 0.04 , sensitivity of 0.94 ± 0.03 , and specificity of 0.79 ± 0.08 (Figure 3). The Clinical-Merged-ML and Clinical-Merged-DL models achieved similar performances, with AUC scores of 0.83 ± 0.01 and 0.83 ± 0.05 , which were both lower than the Clinical-DWI-DL model.

3.4 Robustness of the model on different scanners

We compared the performance of our DWI-DL model on one scanner (GE-Signa HDxt) and all three scanners to assess the robustness of the DL-model on different scanners. We obtained AUC scores of 0.69 ± 0.08 when used with a single scanner and 0.68 ± 0.03 when used with three scanners. We did not detect a statistically significant difference ($p = 0.69$), even though the model performed better on a single scanner than on three scanners.

4 Discussion

In this study, we developed and validated a variety of models for non-invasive preoperative prediction of TDs in patients with RC, based on radiomic features, clinical factors, and a combination of both. Among all the models, the integrated DL-based model using a combination of DWI radiomic features and clinical characteristics was the most effective and achieved promising predictive performance. This approach can serve as a potential preoperative assessment tool to assist clinicians in preoperative stage evaluation and personalized treatment of patients with RC.

Of the 500 included RC patients, 26% presented with TDs, which is slightly higher than the median incidence (21.3%) of TDs in patients with CRC, as previously reported (2). TDs are an important prognostic factor in CRC, as a significantly worse prognosis has been found in patients with TDs, regardless of the sub-staging of the LNs (35). TDs are also an independent risk factor for liver, lung, and peritoneal

TABLE 2 The baseline characteristics of the patients.

Characteristics		TDs-positive	TDs-negative	P
Gender	Male	49(36.8)	136(37.1)	0.96
	Female	84(63.0)	231(62.9)	
Age		65.0(11.0)	64.4(10.5)	0.60
BMI		23.2(2.6)	22.9(3.2)	0.49
CEA (ng/ml)		16.5(40.3)	6.3(11.4)	<0.01
CA 19-9 (IU/ml)		61.7(300.6)	27.1(217.9)	0.16
Distance(cm)		8.8(7.2)	8.9(3.4)	0.81
Tumor length (cm)		4.6(1.4)	4.1(1.3)	<0.01
mrT stage	T1	1(0.8)	8(2.2)	<0.01
	T2	14(10.5)	101(27.5)	
	T3	86(64.7)	228(62.1)	
	T4	32(24.1)	30(8.2)	
mrN stage	N0	2(1.5)	226(61.6)	<0.01
	N1	85(63.9)	101(27.5)	
	N2	46(34.6)	40(10.9)	
CRM	Presence	44(33.1)	53(14.4)	<0.01
	Absence	89(66.9)	314(85.6)	
EVMI	Presence	57(42.9)	68(18.5)	<0.01
	Absence	76(57.1)	299(81.5)	
pT stage	T1	1(0.8)	32(8.7)	<0.01
	T2	13(9.8)	121(33.0)	
	T3	109(81.9)	184(50.1)	
	T4	10(7.5)	30(8.2)	
pN stage	N0	0(0)	272(74.1)	<0.01
	N1a	32(24.1)	38(10.4)	
	N1b	26(19.5)	35(9.5)	
	N1c	35(26.3)	0(0)	
	N2a	27(20.3)	14(3.8)	
	N2b	13(9.8)	8(2.2)	
LI	Presence	29(21.8)	19(5.2)	<0.01
	Absence	104(78.2)	348(94.8)	
PI	Presence	40(30.1)	28(7.6)	<0.01
	Absence	93(69.9)	339(92.4)	
Grade	Well differentiated	47(35.3)	181(49.3)	0.01
	Moderately differentiated	61(45.9)	145(39.5)	
	Poorly/undifferentiated	25(18.8)	41(11.2)	

Unless otherwise indicated, data are the number of patients, with percentages in parentheses. Categorical variables were compared by using the chi-squared test. $P < 0.05$ indicates a statistically significant difference. Continuous variables were expressed as means \pm standard deviations. TD, tumor deposition; CEA, carcinoembryonic antigen; CA19-9, carbohydrate antigen 19-9; mrT stage, tumor stage on magnetic resonance imaging; mrN stage, lymph node stage on magnetic resonance imaging; CRM, circumferential resection margin; EMVI, extramural microvascular invasion; pT stage, pathological tumor stage; pN stage, pathological lymph node stage; LI, lymphovascular invasion; PI, perineural invasion; Grade, pathological tumor histological grade.

TABLE 3 Comparison of areas under the curve for all models.

Models	DL				ML			
	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE
DWI	0.68 ± 0.03	0.674 ± 0.03	0.708 ± 0.07	0.66 ± 0.06	0.64 ± 0.08	0.64 ± 0.06	0.71 ± 0.20	0.58 ± 0.26
HRT2	0.57 ± 0.06	0.63 ± 0.05	0.57 ± 0.18	0.65 ± 0.09	0.62 ± 0.02	0.62 ± 0.02	0.58 ± 0.21	0.66 ± 0.22
Merged (DWI+HRT2)	0.59 ± 0.04	0.63 ± 0.07	0.66 ± 0.10	0.62 ± 0.13	0.69 ± 0.04	0.67 ± 0.04	0.70 ± 0.07	0.66 ± 0.06
Clinical	0.81 ± 0.04	0.76 ± 0.05	0.93 ± 0.05	0.72 ± 0.07	0.81 ± 0.06	0.80 ± 0.02	0.96 ± 0.05	0.63 ± 0.07
Clinical-DWI	0.90 ± 0.04	0.84 ± 0.05	0.94 ± 0.03	0.79 ± 0.08	0.81 ± 0.02	0.77 ± 0.03	0.87 ± 0.04	0.68 ± 0.07
Clinical-HRT2	0.83 ± 0.04	0.75 ± 0.04	0.94 ± 0.04	0.69 ± 0.05	0.79 ± 0.02	0.76 ± 0.02	0.77 ± 0.06	0.75 ± 0.07
Clinical-Merged (DWI+HRT2)	0.83 ± 0.05	0.75 ± 0.06	0.97 ± 0.02	0.67 ± 0.08	0.83 ± 0.01	0.74 ± 0.03	0.86 ± 0.08	0.74 ± 0.07

Bold results indicate better results. The integrated model of clinical factors and diffusion-weighted imaging obtained the best performance.

metastases (36). Moreover, patients with TDs have a higher risk of LN metastasis and lymphovascular and perineural invasion (37). In this study, the proportion of patients with LN metastases and vascular and nerve invasion in the TD+ group (73.7%, 21.8%, and 30.1%, respectively) was also significantly higher than that in the TD− group (25.9%, 5.2%, and 7.6%, respectively), indicating a possible correlation between TDs and LN metastases, neurovascular invasion, and multi-channel tumor metastases, which are also associated with worse prognosis in patients with CRC.

While TDs cannot be reliably assessed preoperatively using traditional imaging techniques that depend on the naked eye, previous studies have shown that they may be predicted using radiomics, which provides implicit information on tumor heterogeneity far beyond the capability of visual inspection. Radiomic models based on US, CT, and MRI have been established for TD prediction. However, functional MRI (e.g., DWI) provides more information on tumor heterogeneity. Therefore, we established radiomic models based on DWI, which demonstrated higher predictive performance than HRT2-only radiomic models.

The Clinical-ML and Clinical-DL models perform similarly in Table 3. This is because there are only 7 clinical features, which is a relatively small number, and both DL and ML work well with such low-dimensional data. We can also observe that the Clinical-DWI-DL

model improves by roughly 9% over the Clinical-DL model, while the Clinical-DWI-ML model barely improves. This could be because deep learning models outperform ML-based models in high-dimensional data situations.

A study on MRI evaluation demonstrated that a joint-modality (HRT2 and DWI) radiomic model achieved higher diagnostic performance than HRT2-only and DWI-only models (38). However, in our study, the joint-modality model did not outperform single-modality models. This result is similar to the findings of Shin et al. who predicted the complete pathological response in RC, and their joint-modality model using features from T2-weighted and DWI images had a classification performance similar to that of the T2-only model (39). We also developed an integrated model that combined radiomic features and clinical characteristics to improve the predictive performance. Among the integrated DL-based models combined with clinical factors, the model utilizing DWI-only radiomic features achieved the highest performance. This may be due to inconsistent baselines and different spatial resolutions between HRT2 and DWI scans of RC, which cannot be reduced by spatial resampling prior to feature extraction.

In addition to investigating the model's performance, we investigated its robustness due to the complexity of clinical data collection. We chose the model with the best performance (DWI-DL) for this investigation because clinical information is independent

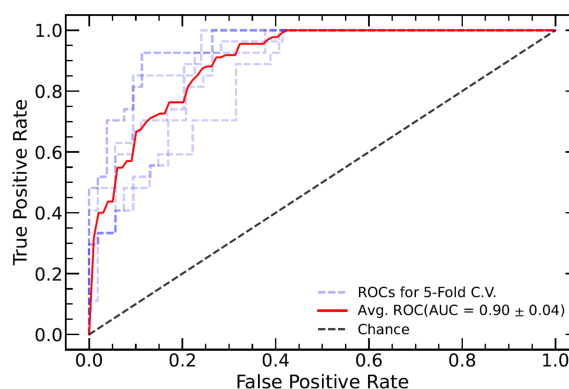


FIGURE 3

ROC curve of the combined model of DWI and clinical factors. The blue dotted lines represent the ROC curve of five-fold cross-validation (CV), and the red line represents the mean ROC curve of the five-fold CV. ROC, receiver operating characteristic; DWI, diffusion-weighted imaging.

of the scanner, allowing us to test the model's robustness across a range of scanners. We found no significant differences in the radiomics model across machines (AUCs of the single- and multi-scanner models: 0.69 ± 0.08 and 0.68 ± 0.03 , respectively, with $P = 0.69$), indicating good robustness of our radiomics model.

This study has some limitations. First, to our knowledge, this is the largest study to date on TD radiomics research, but it is still not large enough to avoid selection bias that compromises the generalization ability of our models. Second, this retrospective study excluded patients who had received neoadjuvant chemotherapy or radiotherapy before surgery, which introduced a further selection bias. Third, this was a single-center study, and the difference in sample sizes between the TD+ and TD- groups was large. Therefore, further prospective, multicenter studies with larger cohorts are warranted to improve prediction outcomes and define the potential standardization of our models.

5 Conclusions

Our integrated model combining clinical variables (tumor markers and MRI reporting status) and MRI radiomic features in a DL model can non-invasively and preoperatively predict TDs in patients with RC. In particular, the model that used DWI and clinical features showed the highest predictive performance. This model can serve as a potential preoperative assessment tool in clinical practice for more effective tumor staging and risk stratification to provide optimal treatment for patients with RC.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

References

- Delattre JF, Cohen R, Henriques J, Falcoz A, Emile JF, Fratte S, et al. Prognostic value of tumor deposits for disease-free survival in patients with stage III colon cancer: A *Post hoc* analysis of the IDEA France phase III trial (PRODIGE-GERCOR). *J Clin Oncol* (2020) 38(15):1702–10. doi: 10.1200/JCO.19.01960
- Lord AC, D'Souza N, Pucher PH, Moran BJ, Abulafi AM, Wotherspoon A, et al. Significance of extranodal tumour deposits in colorectal cancer: A systematic review and meta-analysis. *Eur J Cancer* (2017) 82:92–102. doi: 10.1016/j.ejca.2017.05.027
- Shi M, Zhang H, Yao G, Wu J, Zhu C, Zhang X, et al. The role of tumor deposits in predicting the efficacy of chemotherapy in stage III colon cancer. *Front Oncol* (2020) 10:586603. doi: 10.3389/fonc.2020.586603
- Wong-Chong N, Motl J, Hwang G, Nassif GJ Jr, Albert MR, Monson JRT, et al. Impact of tumor deposits on oncologic outcomes in stage III colon cancer. *Dis Colon Rectum* (2018) 61(9):1043–52. doi: 10.1097/DCR.0000000000001152
- Kaur H, Choi H, You YN, Rauch GM, Jensen CT, Hou P, et al. MR imaging for preoperative evaluation of primary rectal cancer: Practical considerations. *Radiographics* (2012) 32(2):389–409. doi: 10.1148/rg.322115122
- Horvat N, Carlos Tavares Rocha C, Clemente Oliveira B, Petkovska I, Gollub MJ. MRI Of rectal cancer: Tumor staging, imaging techniques, and management. *Radiographics* (2019) 39(2):367–87. doi: 10.1148/rg.2019180114
- Gröne J, Loch FN, Taupitz M, Schmidt C, Kreis ME. Accuracy of various lymph node staging criteria in rectal cancer with magnetic resonance imaging. *J Gastrointest Surg* (2018) 22(1):146–53. doi: 10.1007/s11605-017-3568-x
- Langman G, Patel A, Bowley DM. Size and distribution of lymph nodes in rectal cancer resection specimens. *Dis Colon Rectum* (2015) 58(4):406–14. doi: 10.1097/DCR.0000000000000321
- Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rödel C, Cervantes A, et al. Rectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2017) 28(suppl_4):iv22–40. doi: 10.1093/annonc/mdx224
- Jin Y, Li M, Zhao Y, Huang C, Liu S, Liu S, et al. Computed tomography-based radiomics for preoperative prediction of tumor deposits in rectal cancer. *Front Oncol* (2021) 11:710248. doi: 10.3389/fonc.2021.710248
- Benson AB, Venook AP, Al-Hawary MM, Arain MA, Chen YJ, Ciombor KK, et al. NCCN guidelines insights: Rectal cancer, version 6.2020. *J Natl Compr Canc Netw* (2020) 18(7):806–15. doi: 10.6004/jnccn.2020.0032
- Lord AC, D'Souza N, Shaw A, Rakan Z, Moran B, Abulafi M, et al. MRI-Diagnosed tumor deposits and EMVI status have superior prognostic accuracy to current clinical TNM staging in rectal cancer. *Ann Surg* (2022) 276(2):334–44. doi: 10.1097/SLA.0000000000004499
- You J, Yin J. Performances of whole tumor texture analysis based on MRI: Predicting preoperative T stage of rectal carcinomas. *Front Oncol* (2021) 11:678441. doi: 10.3389/fonc.2021.678441
- Lu H, Yuan Y, Zhou Z, Ma X, Shen F, Xia Y, et al. Assessment of MRI-based radiomics in preoperative T staging of rectal cancer: Comparison between minimum and maximum delineation methods. *BioMed Res Int* (2021) 2021:5566885. doi: 10.1155/2021/5566885

Author contributions

CF, LZ and FZ conceived of the presented idea. MH, JQ, PL, ZY, KS, MW, XW and FL collected the data. LZ, CF, JZ, WL and TS analyzed the data. FL drafted the manuscript. TS and JS improved the quality of English. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by Health Commission of Zhejiang Province (WKJ-ZJ-2210), Science Technology Department of Zhejiang Province (2019C03014), and the National Natural Science Foundation of China (81871403).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

15. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol* (2016) 34(18):2157–64. doi: 10.1200/JCO.2015.65.9128
16. Li C, Yin J. Radiomics based on T2-weighted imaging and apparent diffusion coefficient images for preoperative evaluation of lymph node metastasis in rectal cancer patients. *Front Oncol* (2021) 11:671354. doi: 10.3389/fonc.2021.671354
17. Zhang K, Ren Y, Xu S, Lu W, Xie S, Qu J, et al. A clinical-radiomics model incorporating T2-weighted and diffusion-weighted magnetic resonance images predicts the existence of lymphovascular invasion / perineural invasion in patients with colorectal cancer. *Med Phys* (2021) 48(9):4872–82. doi: 10.1002/mp.15001
18. Albaradei S, Thafar M, Alsaedi A, Van Neste C, Gojobori T, Essack M, et al. Machine learning and deep learning methods that use omics data for metastasis prediction. *Comput Struct Biotechnol J* (2021) 19:5008–18. doi: 10.1016/j.csbj.2021.09.001
19. Liu Z, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang Z, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res* (2017) 23(23):7253–62. doi: 10.1158/1078-0432.CCR-17-1038
20. Horvat N, Veeraraghavan H, Khan M, Blazic I, Zheng J, Capanu M, et al. MR imaging of rectal cancer: Radiomics analysis to assess treatment response after neoadjuvant therapy. *Radiology* (2018) 287(3):833–43. doi: 10.1148/radiol.2018172300
21. Esteve A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* (2017) 542(7639):115–8. doi: 10.1038/nature21056
22. Ragab DA, Sharkas M, Marshall S, Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* (2019) 7:e6201. doi: 10.7717/peerj.6201
23. Kumar Y, Gupta S, Singla R, Hu YC. A systematic review of artificial intelligence techniques in cancer prediction and diagnosis. *Arch Comput Methods Eng* (2022) 29(4):2043–70. doi: 10.1007/s11831-021-09648-w
24. Chen LD, Li W, Xian MF, Zheng X, Lin Y, Liu BX, et al. Preoperative prediction of tumour deposits in rectal cancer by an artificial neural network-based US radiomics model. *Eur Radiol* (2020) 30(4):1969–79. doi: 10.1007/s00330-019-06558-1
25. Yang YS, Feng F, Qiu YJ, Zheng GH, Ge YQ, Wang YT. High-resolution MRI-based radiomics analysis to predict lymph node metastasis and tumor deposits respectively in rectal cancer. *Abdom Radiol (NY)* (2021) 46(3):873–84. doi: 10.1007/s00261-020-02733-x
26. Beets-Tan RGH, Lambregts DMJ, Maas M, Bipat S, Barbaro B, Curvo-Semedo L, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European society of gastrointestinal and abdominal radiology (ESGAR) consensus meeting. *Eur Radiol* (2018) 28(4):1465–75. doi: 10.1007/s00330-017-5026-2
27. Kalisz KR, Enzerra MD, Paspulati RM. MRI Evaluation of the response of rectal cancer to neoadjuvant chemoradiation therapy. *Radiographics* (2019) 39(2):538–56. doi: 10.1148/rg.2019180075
28. Nougaret S, Reinhold C, Mikhael HW, Rouanet P, Bibeau F, Brown G. The use of MR imaging in treatment planning for patients with rectal carcinoma: have you checked the "DISTANCE"? *Radiology* (2013) 268(2):330–44. doi: 10.1148/radiol.13121361
29. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* (2006) 31(3):1116–28. doi: 10.1016/j.neuroimage.2006.01.015
30. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, et al. CE-net: Context encoder network for 2D medical image segmentation. *IEEE Trans Med Imaging* (2019) 38(10):2281–92. doi: 10.1109/TMI.2019.2903562
31. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77(21):e104–e7. doi: 10.1158/0008-5472.CAN-17-0339
32. Lemaitre G, Nogueira F, Aridas CK. Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning. *ArXiv* (2017). doi: 10.48550/arXiv.1609.06570
33. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* (2007) 26(30):5512–28. doi: 10.1002/sim.3148
34. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. *PyTorch: An imperative style, high-performance deep learning library*. *NeurIPS*. Vancouver CANADA: Curran Associates, Inc. (2019).
35. Bai R, Tan Y, Li D, Yang M, Yu L, Yuan Y, et al. Development and validation of a novel prognostic nomogram including tumor deposits could better predict survival for colorectal cancer: A population-based study. *Ann Transl Med* (2021) 9(8):620. doi: 10.21037/atm-20-4728
36. Cohen R, Shi Q, Meyers J, Jin Z, Svrcek M, Fuchs C, et al. Combining tumor deposits with the number of lymph node metastases to improve the prognostic accuracy in stage III colon cancer: A *post hoc* analysis of the CALGB/SWOG 80702 phase III study (Alliance)(☆). *Ann Oncol* (2021) 32(10):1267–75. doi: 10.1016/j.annonc.2021.07.009
37. Nagtegaal ID, Knijn N, Hugen N, Marshall HC, Sugihara K, Tot T, et al. Tumor deposits in colorectal cancer: Improving the value of modern staging—a systematic review and meta-analysis. *J Clin Oncol* (2017) 35(10):1119–27. doi: 10.1200/JCO.2016.68.9091
38. Park SH, Cho SH, Choi SH, Jang JK, Kim MJ, Kim SH, et al. MRI Assessment of complete response to preoperative chemoradiation therapy for rectal cancer: 2020 guide for practice from the Korean society of abdominal radiology. *Korean J Radiol* (2020) 21(7):812–28. doi: 10.3348/kjr.2020.0483
39. Shin J, Seo N, Baek SE, Son NH, Lim JS, Kim NK, et al. MRI Radiomics model predicts pathologic complete response of rectal cancer following chemoradiotherapy. *Radiology* (2022) 303(2):351–8. doi: 10.1148/radiol.211986



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

Aleksandr Shulyak,
National Academy of Medical Sciences of
Ukraine, Ukraine
Andrea Benedetto Galosi,
Marche Polytechnic University, Italy

*CORRESPONDENCE

Nianyu Xue
✉ xuenianyu010133@126.com
Guoyao Wang
✉ wgy9@163.com
Shengmin Zhang
✉ 1261142692@qq.com

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 06 November 2022

ACCEPTED 06 February 2023

PUBLISHED 20 February 2023

CITATION

Xue N, Wang G, Zhang S and Lu Y (2023)
The value of contrast-enhanced
ultrasonography in differential diagnosis of
primary testicular germ cell tumors and
non-germ cell tumors over 50 years old.
Front. Oncol. 13:1090823.
doi: 10.3389/fonc.2023.1090823

COPYRIGHT

© 2023 Xue, Wang, Zhang and Lu. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

The value of contrast-enhanced ultrasonography in differential diagnosis of primary testicular germ cell tumors and non-germ cell tumors over 50 years old

Nianyu Xue^{1*}, Guoyao Wang^{2*}, Shengmin Zhang^{1*} and Yijun Lu³

¹Department of Ultrasonography, Ningbo First Hospital, Ningbo, China, ²Department of Urology, Ningbo First Hospital, Ningbo, Zhejiang, China, ³School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Background: Unlike young and middle-aged patients, seminoma is not common in patients with primary testicular tumors over the age of 50, so it cannot follow the general ideas and norms for diagnosing and treating testicular tumors, and its characteristics need to be considered separately.

Methods: The conventional ultrasonography and contrast-enhanced ultrasonography (CEUS) findings of primary testicular tumors in patients over 50 years old were retrospectively analyzed and compared with the pathological results to compare the diagnostic value of these two methods.

Results: Of the 13 primary testicular tumors, 8 were primary lymphomas. Conventional ultrasound of 13 cases of testicular tumors showed hypoechoic with rich blood flow, and it was difficult to identify the type accurately. The sensitivity, specificity, positive predictive value, negative predictive value and accuracy of conventional ultrasonography in diagnosing non-germ cell tumors (lymphoma and Leydig cell tumor) were 40.0%, 33.3%, 66.7%, 14.3%, and 38.5%, respectively. CEUS findings: 7 of 8 lymphomas showed uniform hyperenhancement. 2 cases of Leydig cell tumors showed uniform high enhancement. 2 cases of seminoma and 1 case of spermatocytic tumor showed heterogeneous enhancement, with necrosis in the interior. The sensitivity, specificity, positive predictive value, negative predictive value and accuracy rate of non-germ cell tumor diagnosis according to the non-necrotic area of CEUS were 90.0%, 100.0%, 100.0%, 75.0% and 92.3%, respectively. Compared with conventional ultrasound, the difference was statistically significant ($P=0.039$).

Conclusions: Primary testicular tumors in patients over 50 years old are mainly lymphoma, and CEUS is significantly different between germ cell tumors and non-germ cell tumors. Compared with conventional ultrasound, CEUS can distinguish testicular germ cell tumors from non-germ cell tumors more accurately. Preoperative ultrasonography is significant for accurate diagnosis and can guide clinical treatment.

KEYWORDS

ultrasonography, contrast-enhanced ultrasonography, testis, germ cell tumors, non-germ cell tumors

Introduction

Testicular tumors account for 1-2% of male malignant tumors (1). The most common malignant tumors in adolescent males are divided into primary tumors and secondary tumors. Due to the existence of primary lesions and systemic manifestations, secondary tumors are relatively easy to diagnose. Although there are many pathological types of testicular tumors, most of them are germ cell tumors (2, 3). Ultrasound, the first-choice imaging examination for scrotal disease (3, 4), is easy to detect masses, but qualitative diagnosis is more difficult. Since most testicular tumors are malignant, seminoma is the most common (5), and needle biopsy can lead to local recurrence. Therefore, a needle biopsy is generally not recommended for testicular tumors. For testicular tumors smaller than 15 mm, the testicles can be preserved in conjunction with intraoperative frozen sections (6, 7). However, this inevitably leads to a longer surgery time. While for large lesions >4 cm, even if lymphoma is suspected, radical orchiectomy is mandatory in most of cases to relieve symptoms and reduce tumor bulk. However, if lymphoma is diagnosed preoperatively by ultrasound, a puncture biopsy can be taken to make a definitive diagnosis, thus saving time for surgery. However, seminoma is uncommon in elderly patients (2), so it cannot follow the general ideas and norms for diagnosing and treating testicular tumors, and its characteristics need to be considered separately. This study retrospectively analyzed the conventional ultrasonography and contrast-enhanced ultrasonography of primary testicular tumors in patients over 50 years old, in order to make an accurate preoperative diagnosis and guide clinical treatment.

Methods

Subjects

This retrospective study was approved by the Ethics Committee of The First Ningbo Hospital (2021RS105). A retrospective analysis of 14 cases of primary testicular tumors confirmed by pathology in our hospital from January 2013 to December 2021. Inclusion criteria: (1) confirmed by histopathology; (2) with complete

routine ultrasound and CEUS data; (3) age > 50 years. Exclusion criteria: (1) incomplete data: lack of histological and pathological results, incomplete ultrasound images; (2) received non-steroidal anti-inflammatory drugs, radiotherapy, chemotherapy and other immunotherapy; (3) Elevated AFP or HCG. The age ranged from 51 to 67 years old, with an average of (59.6 ± 5.3) years old. According to the pathological results, they were divided into non-germ cell tumor group (lymphoma and Leydig cell tumor) and germ cell tumor group (seminoma and spermatogenic tumor).

Equipment and agents

The contrast agent used in CEUS was SonoVue (Bracco SpA, Milan, Italy). The agents were microbubbles of the phospholipids microencapsulated sulfur hexafluoride (SF₆). The microbubbles had an average diameter of 2.5 μ m and pH values of 4.5–7.5. After the SonoVue powder was thoroughly dissolved in 5 mL of normal saline, 2.4 mL of the solution was injected into the bolus through the cubital vein.

The ultrasound devices used included the Aplio500 (TOSHIBA CORPORATION, Tokyo, Japan), LOGIQ E9 GE (General Electric Company, Boston, Massachusetts, USA), EPIQ7 (Philips Electronic N.V., Amsterdam, The Netherlands), EUB-8500 (HITACHI, Tokyo, Japan), and Aixplorer (SuperSonic Imagine, Aix-en-Provence, France). The CEUS function was available on all of these devices. A linear array probe was used (frequency 5.0–12.0 MHz).

Methodology

The conventional ultrasound and CEUS images of the primary testicular tumors were retrospectively analyzed. The parameters of the conventional ultrasound images included the location, number, size, shape, echo, boundary, and blood flow. The peripheral annular blood flow was used as the diagnostic criterion for Leydig cell tumors and the perforating vessel was used as the diagnostic criteria for lymphoma. If the testicular tumor is hypoechoic with abundant blood flow and irregular blood flow distribution, it is diagnosed as germ cell tumor. The findings were interpreted by 2 physicians with

10 years of experience in scrotal ultrasound, and each preliminary diagnosis was made after the physicians reached an agreement. The parameters of CEUS included the enhancement time, enhancement level (high, equal, low, or none), and contrast-agent distribution (uniform or non-uniform). CEUS showing uniform high enhancement was used as the diagnostic criterion for non-germ cell tumor (lymphoma and Leydig cell tumor). CEUS showing heterogeneous enhancement with necrosis in the interior was used as the diagnostic criterion for germ cell tumors of the testis. The findings were interpreted by 2 physicians with 5 years of experience in scrotal ultrasound, and each preliminary diagnosis was made after the physicians reached an agreement. All 4 physicians were blind to the final diagnoses and other imaging information at the time of the interpretation and preliminary diagnoses.

The clinical flow chart for the diagnosis and treatment of testicular tumors is shown in Figure 1.

Statistical analysis

The statistical analysis was performed using the SPSS13.0 software package (Chicago, IL, USA). The count data were analyzed using the paired χ^2 test, and the diagnostic accuracy of the examinations was assessed in 2-by-2 tables. The measurement data were expressed as the mean \pm SD. A P value <0.05 was considered statically significant.

Results

Clinicopathological data

A total of 12 patients (13 masses in total), except for 1 case of bilateral testicular lymphoma, the others were unilateral and single.

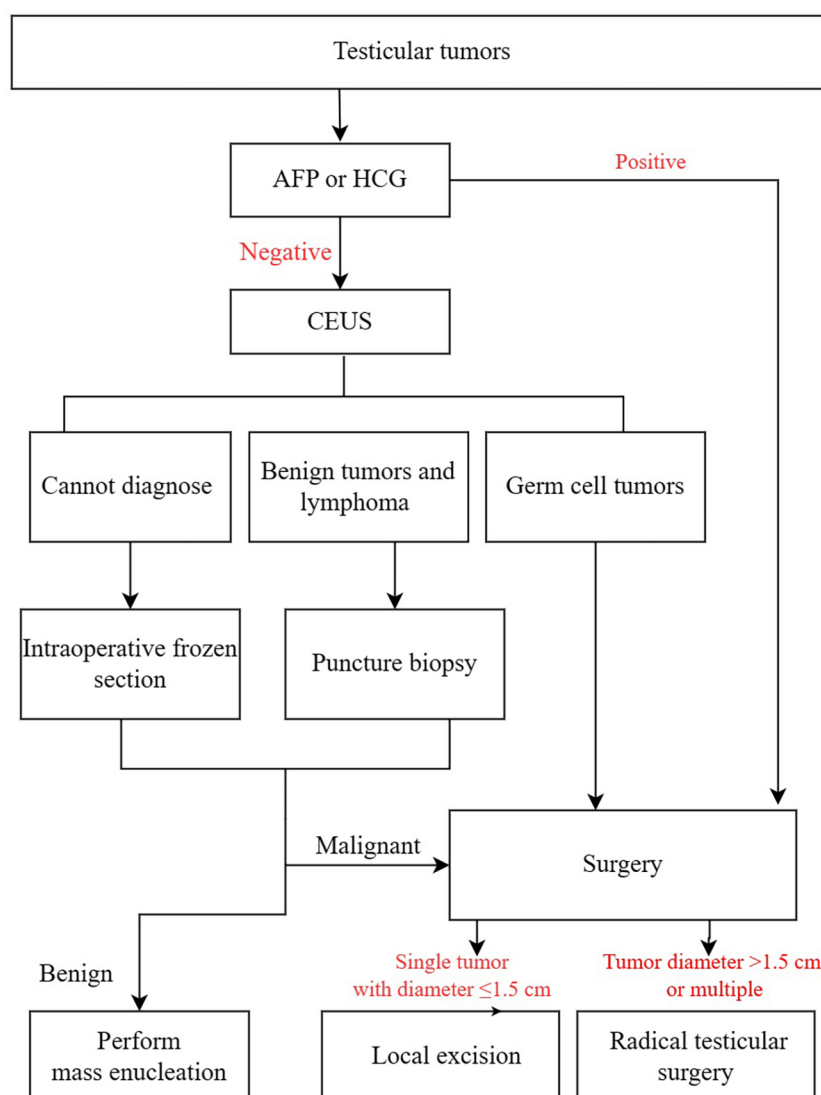


FIGURE 1
Clinical flow chart.

Among the 13 cases of primary testicular tumors confirmed by pathology, 10 were allocated to the non-germ cell tumor group (including 8 cases of primary lymphoma and 2 cases of Leydig cell tumors) and 3 were allocated to the germ cell tumor group (including 2 cases of seminomas, and 1 case of spermatocytic tumor). The lesions occurred in the left testis in 8 patients and in the right testis in 5 patients.

Except for 2 cases of Leydig cell tumors found by ultrasonography, the others were all due to scrotal enlargement (1 case of bilateral testicular lymphoma with scrotal pain, the rest were painless). 1 case of bilateral lymphoma had elevated follicle-stimulating hormone and luteinizing hormone, and no other related abnormalities were found.

Findings of conventional ultrasound and CEUS

The characteristics of conventional ultrasound and CEUS in testicular germ cell tumors and non-germ cell tumors over 50 years old are shown in Table 1. All 13 showed abundant blood flow, including 3 lymphomas, 1 spermatocytic tumor, and 1 seminoma with perforating vessels (Figures 2, 3), and 1 Leydig cell tumor with peripheral circular blood flow (Figure 4), other blood flow is irregular.

The characteristics of conventional ultrasound in 13 cases were similar, and it was difficult to identify their pathological types accurately. The sensitivity, specificity, positive predictive value, negative predictive value and accuracy rate of conventional ultrasound in the diagnosis of non-germ cell tumors were respectively 40.0%, 33.3%, 66.7%, 14.3%, and 38.5%.

CEUS: 8 cases of lymphoma showed fast forward and fast regression, 7 cases showed uniform high enhancement, no obvious necrosis area (Figure 2), 1 case of lymphoma showed obvious necrosis area; 2 cases of seminoma showed high enhancement around the periphery, with necrosis in the interior, showing fast forward and fast regression; 1 case of spermatocytic tumor showed uneven, sparse, low enhancement, fast forward and equal regression (Figure 3); 2 cases of Leydig cell tumors showed uniform high enhancement, no obvious necrotic area, and showed rapid progress and slow regression (Figure 4). Spermatocytic tumor was the only tumor with the low enhancement of all tumors. Lymphoma and Leydig cell tumor showed uniform high enhancement, lymphoma fast forward and fast regression, Leydig cell tumor fast forward and slow regression. The sensitivity, specificity, positive predictive value, negative predictive value and accuracy rate of non-germ cell tumor diagnosis according to the non-necrotic area of CEUS were 90.0%, 100.0%, 100.0%, 75.0% and 92.3%, respectively. Compared with conventional ultrasound, the difference was statistically significant ($P=0.039$).

TABLE 1 Conventional ultrasound and CEUS features of testicular germ cell tumors and non-germ cell tumors over 50 years old.

Conventional ultrasound and CEUS	germ cell tumors group (n=3)	non-germ cell tumors group (n=10)
Age (years), mean \pm SD	58.0 \pm 7.5	60.1 \pm 4.8
Lesion size (cm), mean \pm SD	5.1 \pm 2.4	4.2 \pm 2.7
Boundary, n		
Clear	2	10
Unclear	1	0
Echoes, n		
Hypoechoic	3	10
Shape, n		
Round or oval	2	7
Irregular	2	3
Echo distribution, n		
Homogenous	0	4
Heterogenous	3	6
Blood flow, n		
Perforating vessels	2	3
Peripheral annular	0	1
Irregular	1	6
Enhancement, n		
Uniform hyperenhancement	0	1
Heterogeneous enhancement	3	9

SD, standard deviation; CEUS, contrast-enhanced ultrasonography.

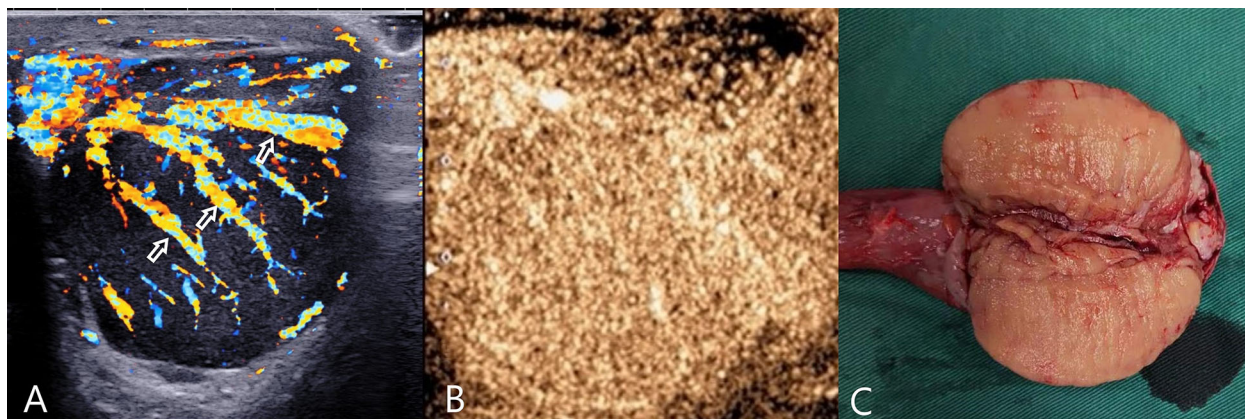


FIGURE 2

Testicular lymphoma. (A) CDFI: hypoechoic mass in the left testis, abundant blood flow, with multiple perforating vessels (arrow); (B) CEUS: CEUS showed uniform and high enhancement, and no obvious necrotic area was found. (C) Macro-section: Mass occupying the entire testicle; CDFI, Color Doppler Flow Imaging; CEUS, contrast-enhanced ultrasonography.

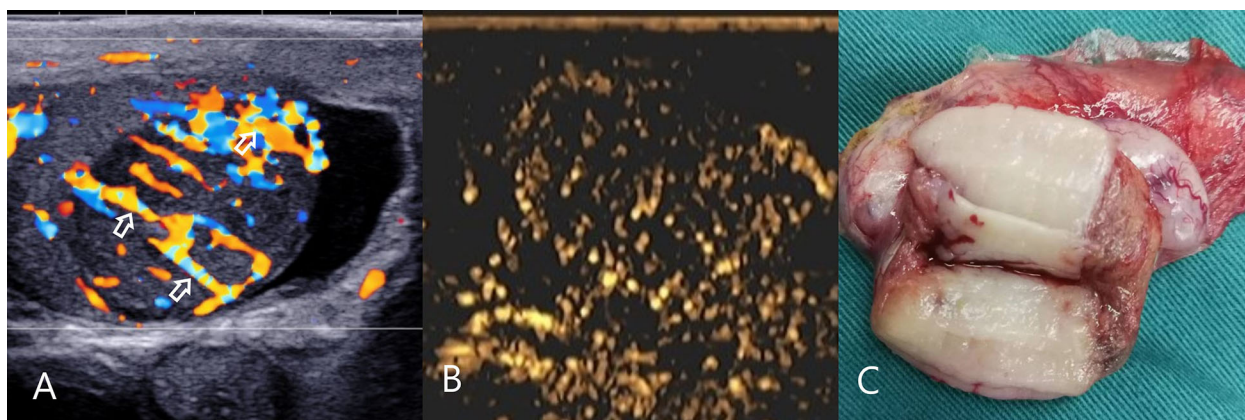


FIGURE 3

Testicular spermatocytic tumor. (A) CDFI: hypoechoic mass in the right testis, abundant blood flow and multiple perforating vessels (arrow), consistent with the ultrasound findings of lymphoma. (B) CEUS: The CEUS showed heterogeneous sparse and low enhancement, which was significantly different from that of lymphoma. (C) Macro-section: Mass occupies most of the testicle; CDFI, Color Doppler Flow Imaging; CEUS, contrast-enhanced ultrasonography.

Discussion

The most common testicular tumor in adolescents is seminoma (5), so the choice of radical orchiectomy is most feasible. However, seminoma is very rare in patients over 50 years old. From this study, lymphoma is the main one. Since lymphoma is a combination treatment requiring surgery, chemotherapy and radiotherapy, intraoperative frozen sections are more difficult to diagnose, so preoperative puncture biopsy for immunohistochemistry can clarify the diagnosis (8). Since Leydig cell tumors are mostly benign, enucleation is the first choice (9, 10). Pathological diagnosis of Leydig cell tumors by preoperative biopsy can reduce the operation time and protect testicular tissue. Therefore, the correct diagnosis is of great significance to the patient and clinical, which also puts forward

higher requirements for ultrasound diagnosis, not just the diagnosis of malignancy.

Relying on the history of the primary tumor and systemic manifestations, secondary testicular tumors are relatively easy to diagnose. Primary testicular tumors are more challenging to diagnose. This study shows that most of the primary testicular tumors are large. This is due to the relatively fast growth rate of lymphoma and embryonal carcinoma. In the early stage of the tumor, the tumor is small and generally has no clinical symptoms. Only when the tumor grows to a certain extent clinical symptoms are detected. Painless enlargement of the scrotum is the most common symptom in most patients (only Leydig cell tumors are occasionally found in the scrotum due to inguinal hernia), but this is also a common feature of testicular tumors and is not specific.

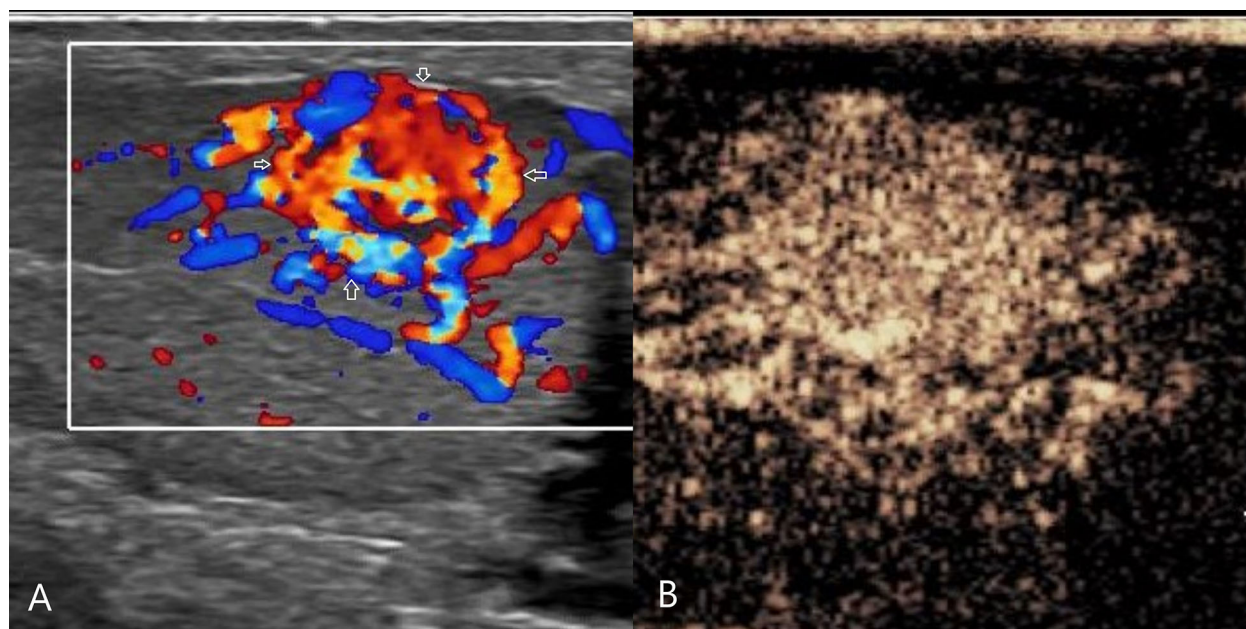


FIGURE 4

Leydig cell tumor. (A) CDFI: hypoechoic mass in the left testis, abundant blood flow, with annular blood flow around (arrow). (B) CEUS: uniform and high enhancement, and no obvious necrotic area was found. CDFI, Color Doppler Flow Imaging; CEUS, contrast-enhanced ultrasonography.

This study showed that 61.5% were lymphoma, and there were only 2 cases of seminoma. Therefore, patients over 50 years old with painless scrotal enlargement should first consider lymphoma.

This study showed that primary testicular tumors had very similar gray-scale ultrasonographic appearances (most of them showed clear borders, round-like hypoechoic), and it was difficult to distinguish pathological types. Color Doppler blood flow has a certain value in diagnosing testicular tumors, and perforating vessels have a certain value in diagnosing lymphoma. The reason for the formation of perforating vessels may be: lymphoma is a disease mainly caused by single-cell proliferation, and the lesions originate in the interstitium of the testis, so the original vascular anatomy in the testis may not be affected, which is different from benign tumors on vascular compression, malignant tumors present differently to vascular compression and erosion. This study showed that perforating vessels appeared in 3 lymphoma masses, and the diagnosis was based on this. The remaining 5 lymphomas only showed abundant blood flow without perforating vessels. 1 case of spermatocytic tumor and 1 case of seminoma had perforating vessels, which were misdiagnosed as lymphoma by conventional ultrasonography. These perforating vessels may be elongated feeding vessels of the tumor. 1 case of Leydig cell tumor showed abundant blood flow with annular blood flow around it. This sign is not seen in other testicular tumors, so peripheral annular blood flow can be used as a specific sign of Leydig cell tumors (to be further confirmed in large sample studies).

In this study, CEUS can accurately distinguish germ cell tumors from non-germ cell tumors (lymphoma and stromal cell tumor) based on the presence or absence of necrosis, with an accuracy rate of 92.3%. Seminoma over the age of 50 is relatively rare. This study showed that the 2 cases of seminoma showed fast forward and fast regression,

heterogeneous high enhancement, and necrotic areas. Combined with negative tumor markers, a diagnosis can be made. Spermatocytic tumor was the only tumor with low enhancement on CEUS of all tumors in this study, which can be diagnosed in combination with negative tumor markers. In the previous WHO classification, spermatocytic tumor was regarded as a subtype of seminoma, called spermatocytic seminoma (11). The 2016 WHO classification separates spermatocytic tumor from seminoma. Although the cytological appearance of these two tumors are similar (12), the CEUS appearance is significantly different. The CEUS of lymphomas all showed fast forward and fast backward, except for 1 case with necrosis. The others showed uniform hyperenhancement. Therefore, in patients over 50 years old, a painless testicular mass, perforating vessels, uniform high enhancement on CEUS, and negative tumor markers can be diagnosed as lymphoma. Leydig cell tumor of the testis generally has a small mass. The CEUS shows uniform high enhancement, fast forward and slow regression, which is different from lymphoma, and can make a clear diagnosis. Enucleation of the mass was performed to avoid the removal of the testis. Pathological diagnosis of Leydig cell tumors and lymphomas by preoperative biopsy can reduce the operation time and protect testicular tissue.

The present study had some limitations. First, this study is a retrospective analysis. The incidence of testicular tumors is low, the sample size is limited, and further research with large samples and multiple centers is needed. Second, limited by the retrospective design and the difference in match models, we did not perform a CEUS-based quantitative analysis. Third, because of the extremely high accuracy of the diagnosis of testicular tumors by the professional andrology sonographers in our center, they have won the trust of andrologists. Most testicular tumors have not undergone magnetic resonance examination, so there is no other image for comparative analysis.

In conclusion, the primary testicular tumors in patients over 50 years old are mainly lymphoma, and CEUS is significantly different between germ cell tumors and non-germ cell tumors. Compared with conventional ultrasound, CEUS can distinguish testicular germ cell tumors from non-germ cell tumors more accurately. Preoperative ultrasonography is of great significance for a clear diagnosis, which can guide clinical treatment and avoid unnecessary orchiectomy.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of The First Ningbo Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Author Contributions Statement: All authors wrote and designed the main manuscript text, and reviewed the manuscript.

NX is responsible for data collection and compilation. NX, GW and SZ are responsible for the analysis and interpretation of data and the final approval of manuscript. YL is responsible for the revision and guidance of language manuscripts. All authors contributed to the article and approved the submitted version.

Acknowledgments

Thanks to Professor Haocheng Lin of Peking University Third Hospital for his guidance and revision of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Albers P, Albrecht W, Algaba F, Bokemeyer C, Cohn-Cedermark G, Fizazi K, et al. EAU guidelines on testicular cancer. *Eur Urol* (2016) 70:8–9.
2. Ambekar A, Rao V, Pai SA, Bindhu MR, Midha D, Kaushal S, et al. Grossing and reporting of testicular tumor specimens: An evidence-based approach. *Indian J Cancer* (2020) 57(1):7–12. doi: 10.4103/ijc.IJC_1072_19
3. Cassell A, Jalloh M, Ndoye M, Yunusa B, Mbodji M, Diallo A, et al. Review of testicular tumor: Diagnostic approach and management outcome in Africa. *Res Rep Urol* (2020) 12:35–42. doi: 10.2147/RRU.S242398
4. Baird DC, Meyers GJ, Hu JS. Testicular cancer: Diagnosis and treatment. *Am Fam Physician* (2018) 97(4):261–8.
5. Stephenson A, Eggener SE, Bass EB, et al. Diagnosis and treatment of early stage testicular cancer: AUA guideline. *J Urol* (2019) 20:10–97. doi: 10.1097/JU.0000000000000318
6. Galosi AB, Fulvi P, Fabiani A, Servi L, Filosa A, Leone L, et al. Testicular sparing surgery in small testis masses: A multinstitutional experience. *Arch Ital Urol Androl* (2016) 88(4):320–4. doi: 10.4081/aiua.2016.4.320
7. Sbröllini G, Mazzaferro D, Ettamimi A, Montironi R, Cordari M, Maselli G, et al. Diagnostic-therapeutic pathway for small lesions of the testis. *Arch Ital Urol Androl* (2014) 86(4):397–9. doi: 10.4081/aiua.2014.4.397
8. Xu H, Yao F. Primary testicular lymphoma: A SEER analysis of 1,169 cases. *Oncol Lett* (2019) 17(3):3113–24. doi: 10.3892/ol.2019.9953
9. Fankhauser CD, Grogg JB, Hayoz S, Wettstein MS, Dieckmann KP, Sulser T, et al. Risk factors and treatment outcomes of 1,375 patients with testicular leydig cell tumors: Analysis of published case series data. *J Urol* (2020) 203(5):949–56. doi: 10.1097/JU.0000000000000705
10. Dell'Atti L, Fabiani A, Palagonia E, Edoardo A, MP P, Scarcella S, et al. Unusual clinical scenarios in urology and andrology. *Arch Ital Urol Androl* (2021) 93(1):120–6. doi: 10.4081/aiua.2021.1.120
11. Qureshi MB, Uddin N, Ahmad Z, Fatima S. Spermatocytic tumor of testis in a young Male. *J Coll Physicians Surg Pak* (2021) 30(6):722–4. doi: 10.29271/jcpsp.2021.06.722
12. Hao ML, Li CH. Spermatocytic tumor: A rare case report. *World J Clin cases* (2021) 9(35):11115–21. doi: 10.12998/wjcc.v9.i35.11115



OPEN ACCESS

EDITED BY

Mitchell Albert,
Lakehead University, Canada

REVIEWED BY

Federica Pediconi,
Sapienza University of Rome, Italy
Min Sun Bae,
Inha University, Republic of Korea

*CORRESPONDENCE

Denis Le Bihan
✉ denis.lebihan@gmail.com

SPECIALTY SECTION

This article was submitted to
Breast Cancer,
a section of the journal
Frontiers in Oncology

RECEIVED 13 July 2022

ACCEPTED 10 January 2023

PUBLISHED 21 February 2023

CITATION

lima M and Le Bihan D (2023) The road to
breast cancer screening with diffusion MRI.
Front. Oncol. 13:993540.
doi: 10.3389/fonc.2023.993540

COPYRIGHT

© 2023 lima and Le Bihan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The road to breast cancer screening with diffusion MRI

Mami lima^{1,2} and Denis Le Bihan^{3*}

¹Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, ²Department of Clinical Innovative Medicine, Institute for Advancement of Clinical and Translational Science, Kyoto University Hospital, Kyoto, Japan, ³NeuroSpin, Joliot Institute, Department of Fundamental Research, Commissariat à l'Energie Atomique (CEA)-Saclay, Gif-sur-Yvette, France

Breast cancer is the leading cause of cancer in women with a huge medical, social and economic impact. Mammography (MMG) has been the gold standard method until now because it is relatively inexpensive and widely available. However, MMG suffers from certain limitations, such as exposure to X-rays and difficulty of interpretation in dense breasts. Among other imaging methods, MRI has clearly the highest sensitivity and specificity, and breast MRI is the gold standard for the investigation and management of suspicious lesions revealed by MMG. Despite this performance, MRI, which does not rely on X-rays, is not used for screening except for a well-defined category of women at risk, because of its high cost and limited availability. In addition, the standard approach to breast MRI relies on Dynamic Contrast Enhanced (DCE) MRI with the injection of Gadolinium based contrast agents (GBCA), which have their own contraindications and can lead to deposit of gadolinium in tissues, including the brain, when examinations are repeated. On the other hand, diffusion MRI of breast, which provides information on tissue microstructure and tumor perfusion without the use of contrast agents, has been shown to offer higher specificity than DCE MRI with similar sensitivity, superior to MMG. Diffusion MRI thus appears to be a promising alternative approach to breast cancer screening, with the primary goal of eliminating with a very high probability the existence of a life-threatening lesion. To achieve this goal, it is first necessary to standardize the protocols for acquisition and analysis of diffusion MRI data, which have been found to vary largely in the literature. Second, the accessibility and cost-effectiveness of MRI examinations must be significantly improved, which may become possible with the development of dedicated low-field MRI units for breast cancer screening. In this article, we will first review the principles and current status of diffusion MRI, comparing its clinical performance with MMG and DCE MRI. We will then look at how breast diffusion MRI could be implemented and standardized to optimize accuracy of results. Finally, we will discuss how a dedicated, low-cost prototype of breast MRI system could be implemented and introduced to the healthcare market.

KEYWORDS

breast cancer, screening, breast imaging, diffusion MRI, standardization, screening techniques, cost, dedicated units

Highlights

Screening has been shown as an effective method to improve the outcome of breast cancer, the leading cause of cancer in women. Mammography is the preferred method due to its low cost and favorable benefit/risk ratio. However, mammography has some limitations, such as exposure to X-rays, difficulty of interpretation in dense breasts, and overdiagnosis. Among other imaging methods, MRI has clearly the highest sensitivity and specificity. Still, MRI is mainly used to manage suspicious lesions revealed by mammography and not for screening, except for a category of well-defined women at risk, due to a high cost and a limited availability. While the standard breast MRI approach relies on the injection of contrast agents, which have their own contraindications, diffusion MRI which delivers information on tissue microstructure and tumor perfusion without the need for contrast agents, has been shown to provide a similar specificity and sensitivity, emerging as a promising alternative approach to breast cancer screening. To achieve this goal, it is necessary to standardize protocols for acquisition and analysis of diffusion MRI data. Second, the accessibility and cost-effectiveness of MRI examinations need to improve significantly, which may become possible with the development of dedicated breast, low-cost units for breast cancer screening.

1 Introduction

With the advent of widespread breast cancer screening by mammography (MMG) in the early to mid-1980s, detection of breast lesions has increased worldwide, and breast cancer is no longer a fatal disease when diagnosed and treated early. Approximately 60% of cancers diagnosed early have a 5-year survival of 99% after treatment and 31% have a 5-year survival of 85% (1). Breast cancer screening has therefore been shown to be an effective method of improving prognosis. In the absence of a reliable blood test, imaging is the primary approach available for screening. MMG has been the reference method until now because it is relatively inexpensive, widely available and has a favorable benefit/risk ratio with good sensitivity and specificity. Nevertheless, MMG suffers from certain limitations, such as exposure to X-rays given the recent discovery that breast tissue is more sensitive to the effects of radiation than most organs. In addition, with MMG, it is often not possible to predict on mammograms whether lesions are malignant, requiring active treatment, or not, so additional investigations must be performed, especially in dense breasts.

Of particular concern is the relatively high rate of overdiagnosis. Recent immunohistochemical studies have revealed that benign proliferative breast disease, most high-grade ductal carcinoma in situ (DCIS), and invasive carcinoma develop along distinct pathways, in contrast to colonic adenoma-carcinoma, which evolves along a single line (2). These findings suggest that different treatment approaches should be offered depending on the nature of the lesion, including therapeutic abstention for benign lesions. For example, while DCIS lesions often do not become invasive, patients diagnosed with DCIS are generally treated as if they were going to have invasive carcinoma. The rate of “overdiagnosis” is estimated to be between 21 and 66% (3). The social, ethical, and economic consequences of such

management of DCIS lesions are enormous: more than 40% of women with DCIS undergo mastectomies, at a rate of some 10,000 per year, so much so that DCIS could be called a “mammographic disease” (4). Clearly, new approaches must be sought to better predict the grade and outcome of diagnosed breast lesions and to reduce burdensome, costly, and potentially unnecessary surgical procedures, such as mastectomy or axillary lymph node excision, whose morbidity is not negligible. It would also reduce surgical scars that could lead to pseudo-lesions on subsequent imaging. Conversely, the sensitivity of MMG for early detection of cancer in breast cancer screening is only 33% (40% for ultrasound) in patients with a high familial risk for breast cancer (lifetime risk $\geq 20\%$), missing some prognostically important diseases (5). Borderline lesions with uncertain malignant potential at biopsy [histologically classified as “B3”, (6)] most often result in a benign end result. However, these lesions are sometimes associated with the simultaneous presence of a malignant tumor with an enhancement rate of between 10 and 35%, and may also act as a risk factor or precursor to malignancy (7, 8). It is therefore necessary to obtain a more accurate classification of lesions at the time of initial diagnosis in order to personalize the therapeutic approach, avoid unnecessary procedures and reduce costs and social burden. With MMG, it is possible to suspect high-grade lesions from the morphology of microcalcifications, but grading is still difficult, with sparse biopsy sampling, because high-grade and low-grade components can coexist in the same patient or even in the same duct. Indeed, MMG may tend to detect slow-growing cancers.

Recently, breast MRI has been successfully introduced in the management of breast cancer. For example, in DCIS, the sensitivity of MRI for accurate assessment of the extent of DCIS is as high as 89%, much higher than MMG, tomosynthesis, or ultrasound (9). Increasing evidence suggests that, overall, breast MRI may be more sensitive, especially for the diagnosis of high-grade DCIS. Breast MRI is often performed by injection of gadolinium-based contrast agents (GBCA), but more recently, diffusion MRI, a completely noninvasive approach that is highly sensitive to changes in tissue microstructure, has been introduced for cancer imaging. Diffusion MRI has both very high sensitivity and specificity for the detection of breast malignancy (10). Diffusion MRI has been successfully used to differentiate between benign and malignant lesions of the breast, as well as tumor extension.

Yet MRI is exceptionally used for breast cancer screening, although supplemental MRI screening in women with extremely dense breast tissue and normal results on MMG has been recommended, as the addition of MRI leads to significantly fewer interval cancers than MMG alone during a 2-year screening period (11). Still, the main problem with breast MRI is that examinations are today performed using expensive general purpose MRI scanners. MRI is therefore performed as a second-line procedure, which adds to the cost of other imaging modalities (MMG and ultrasound), or in specific populations of women. In addition, there are concerns about the side effects of GBCA when performing dynamic contrast-enhanced (DCE) breast MRI (12). Blood tests may become available to screen for certain breast cancers, but they remain largely non-specific today with many false positives or negatives, and imaging will always remain mandatory to localize lesions and personalize treatment. If a dedicated, small-scale, inexpensive breast MRI scanner can be made available, it could be envisaged that one day

MRI could be used as a screening imaging modality, instead of MMG, at least for a larger number of women at moderate to high risk based on personal history, genetic predisposition, or positivity to blood screening tests when these tests become reliable. This view was enthusiastically supported by an international (EU, USA, Asia) committee of breast imaging experts appointed by the European Society of Breast Imaging (EUSOBI) under the chairmanship of Profs. Denis Le Bihan and Julia Camps-Herrero (13). Breast cancer screening represents a huge market. In the United States alone, more than 60 million women over the age of 40 are responsible for 40 million mammograms per year, which corresponds to 65% of the population concerned (14). In contrast, MRI (using standard whole-body systems) accounted for only 0.4% of women aged 25–64 years in 2017 (15).

In this article, we will first review the principles and current status of diffusion MRI of the breast, and evaluate its clinical performance compared with MMG and DCE MRI. We will then discuss how diffusion MRI of the breast could be implemented and standardized to optimize accuracy of results. Finally, we will discuss how a dedicated, low-cost prototype breast MRI system could be implemented and introduced to the healthcare market.

2 Current place of breast MRI in the global management of breast cancer

Breast MRI has been widely available after the introduction of the use of contrast agents (16). Almost all types of breast cancer show detectable patterns of neovascularization with GBCA, which can readily extravasate into the extravascular and extracellular space (17). Thus, the likelihood of breast cancer can be considered extremely low in the absence of contrast enhancement. In practice, contrast-enhanced T1-weighted MRI is the gold standard. Many malignant breast lesions show maximal contrast enhancement in the early phase after injection, with GBCA being removed from the tissue in the late stage (Figure 1). Conversely, benign lesions and normal fibroglandular tissue usually show maximal enhancement in the late stage and of lower amplitude than in malignant lesions, allowing differentiation of these lesions (19). Given the high sensitivity for detection of breast cancer compared to other modalities such as MMG and ultrasound, breast MRI is also used for preoperative evaluation and tumor staging prior to treatment planning, monitoring tumor response to neoadjuvant therapies, to sort scars from recurrences, or in the presence of implants (20, 21).

However, despite its good clinical performance, MRI is usually performed in second intention. Breast cancer screening by MRI is therefore reserved for women with a moderate to high risk of breast cancer (personal history, genetic predisposition, follow-up after breast conserving surgery or contralateral breast screening, mediastinal irradiation, as for Hodgkin's disease, suspicion of specific lesions, such as atypical ductal hyperplasia (ADH), atypical lobular hyperplasia (ALH), and lobular carcinoma *in situ* (LCIS) (22, 23). A first problem is the cost of MRI scans. Efforts have been made to shorten their duration [shortened DCE protocols, (24)]. However, there are also questions regarding GBCA-related side effects. The primary concern regarding nephrogenic systemic fibrosis (NSF) has almost disappeared, at least for DCE after assessment of renal

function (25, 26), as it occurred only in patients with impaired renal dysfunction, and only seven of the 639 cases of patients with biopsy-confirmed NSF to date were discovered after 2008 (with the avoidance of the use of double and triple doses of GBCA that could trigger NSF) (27). The second concern is related to gadolinium retention in tissues, particularly the brain, after repeated exposure to GBCA (28). This risk is particularly important when considering the repeated annual injection of GBCA that would be required for screening (29). Various new approaches are being investigated to mitigate this risk, such as reducing the dose of GBCA. In a recent study, all breast cancers in 41 consecutive women with biopsy-proven breast cancer were detected as small as 0.4 cm with half (0.05 mmol/kg) a dose of gadobutrol on 3T DCE breast MRI (30).

Thus, there is a growing trend toward the use of new approaches based on unenhanced breast MRI for cancer detection (29). Although no consensus has yet been reached, these approaches could open up breast cancer screening to women at intermediate or even low risk for breast cancer. Given its high potential, diffusion MRI would be the obvious candidate for such an approach.

3 Breast diffusion MRI

3.1 Principles

3.1.1 Diffusion-weighted imaging, DWI, and the apparent diffusion coefficient

While the concept of diffusion MRI emerged in the mid-1980s, diffusion MRI has become a mainstay of modern clinical imaging. Diffusion MRI is both a powerful method and concept because diffusing water molecules provide unique information about the microscopic architecture of tissues. Water diffusion is significantly diminished in most malignant tissues, and diffusion MRI, which requires no tracer injection, is rapidly becoming the modality of choice for detecting, characterizing, or even classifying malignant lesions, especially in the breast (31). Diffusion MRI is deeply rooted in the concept that, during their diffusion-induced movements, molecules probe the structure of tissues at a *microscopic* scale, well beyond the usual *millimeter* resolution of images. During typical diffusion imaging times of about 50–100 ms, water molecules move through tissues on average over distances of about 1–15 μm , bouncing off, passing through, or interacting with many tissue components, such as cell membranes, fibers, or macromolecules. Due to the tortuous movement of water molecules around these obstacles ("hindered" diffusion), the actual diffusion distance is reduced compared to free water. Therefore, non-invasive observation of water diffusion-induced displacement distributions *in vivo* provides unique clues to the fine structural features and geometric organization of cells in tissues, as well as to changes in these features as a function of physiological or pathological states.

MRI signals can be sensitized to diffusion by applying a pair of sharp magnetic field gradient pulses, the duration and separation of which can be adjusted to achieve a specific level of diffusion sensitization defined as the "b-value." By acquiring data with different gradient pulse amplitudes, images with different degrees of diffusion sensitivity are obtained. The overall effect of diffusion in the presence of these gradient pulses is a signal attenuation and the MRI

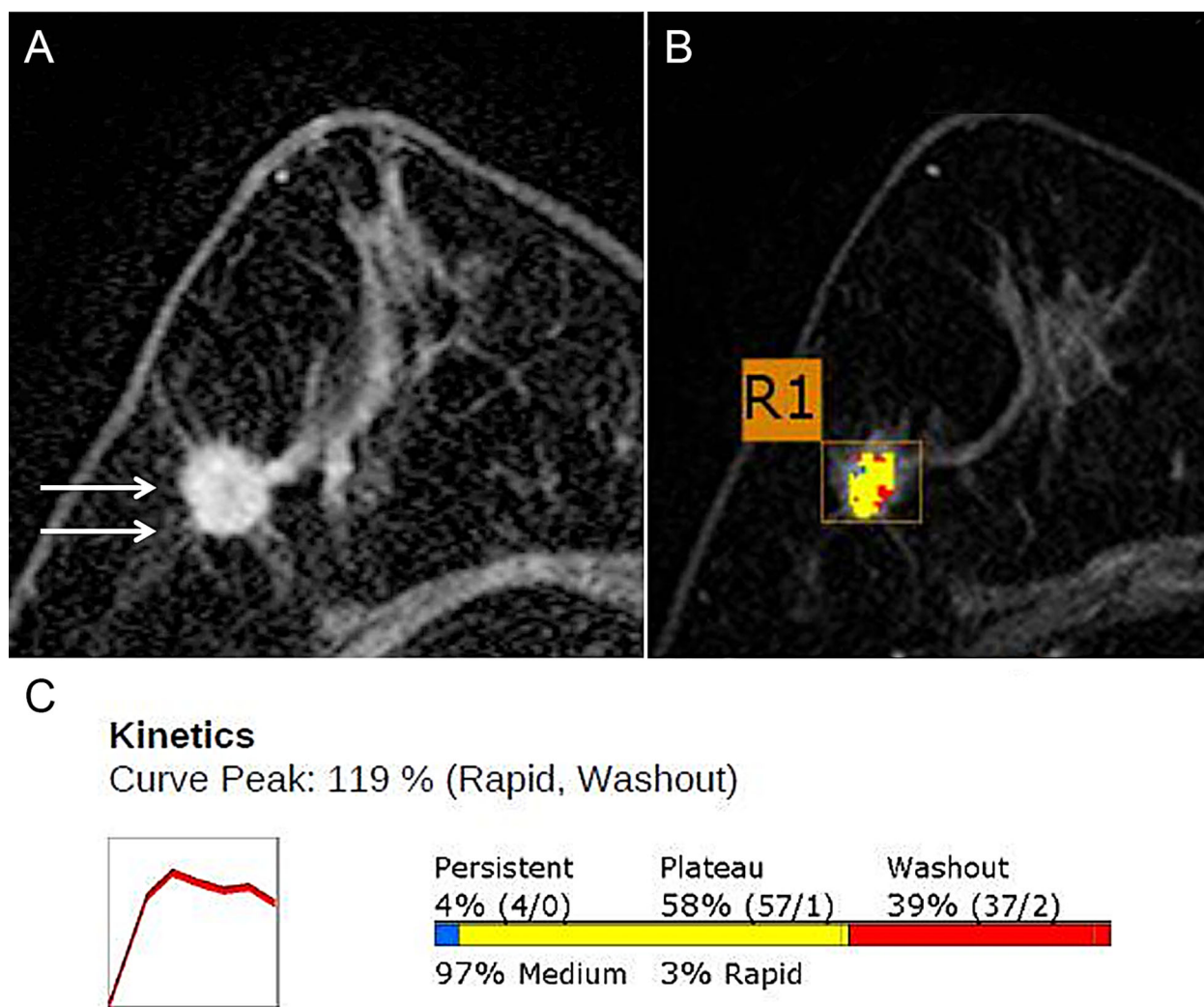


FIGURE 1

DCE-MRI in a 71-year-old Woman with grade 2 invasive ductal carcinoma in the right breast. (A) Axial contrast-enhanced T1w MRI image shows a 15-mm irregular mass (arrows). (B) Color axial maximum-intensity-projection MR image overlaid over the R1 breast mass. A computer-aided detection (CAD) algorithm displays areas in red, yellow, and blue indicating rapid washout-type delayed enhancement, plateau-type delayed enhancement, and persistent-type delayed enhancement patterns, respectively. (C) Graph of the contrast agent uptake shows a rapid initial enhancement and a rapid washout-type curve. The initial peak enhancement value was 119%. With respect to the delayed phase enhancement, 39% of the mass showed washout, 4% of the mass showed a persistent-type curve, and 58% showed a plateau-type curve. Adapted from (18).

signal becomes “diffusion weighted”, hence the term “Diffusion Weighted Imaging” (DWI). The signal attenuation is more pronounced when large values of b are used and when diffusion is fast (because molecules diffuse over larger distances) (Figure 2). It is important to note that only the displacement (diffusion) component in the direction of the gradient pulses is detectable, but the diffusion can be anisotropic.

In DWI, *qualitative* contrast depends not only on diffusion, but also on other MRI parameters, such as T1 and T2 water relaxation times, which can lead to well-known artifacts, such as the “T2-shine-through” effect, as high T2 signal lesions (e.g., necrosis, cysts) can retain a relatively high signal level at high b values. Therefore, these images are often combined numerically to determine a *quantitative* estimate of the diffusion process in each image location, through an *Apparent Diffusion Coefficient* (ADC), “apparent” because diffusion is impeded by many processes (33):

$$ADC = \ln[S(b_0) - S(b_1)] / (b_1 - b_0) \quad (1)$$

where $S(b_0)$ and $S(b_1)$ are the signals (in a voxel or region of interest, ROI) acquired at the b values b_0 and b_1 , respectively. This simple ADC is an incredibly robust and powerful parameter, which has been widely used in all clinical applications of diffusion MRI since its inception (34). The optimal value of b_1 that provides the best contrast-to-noise ratio in breast tissue, i.e., sufficient attenuation of the signal by scattering while maintaining a sufficient signal level is about 800s/mm² (13).

3.1.2 Perfusion and IntraVoxel incoherent motion

Beyond molecular diffusion, blood microcirculation in capillary networks (perfusion) also contributes to the diffusion MRI signal. Indeed, the flow of blood water in pseudo-randomly oriented capillaries (at the voxel level) mimics a random walk (“pseudo-

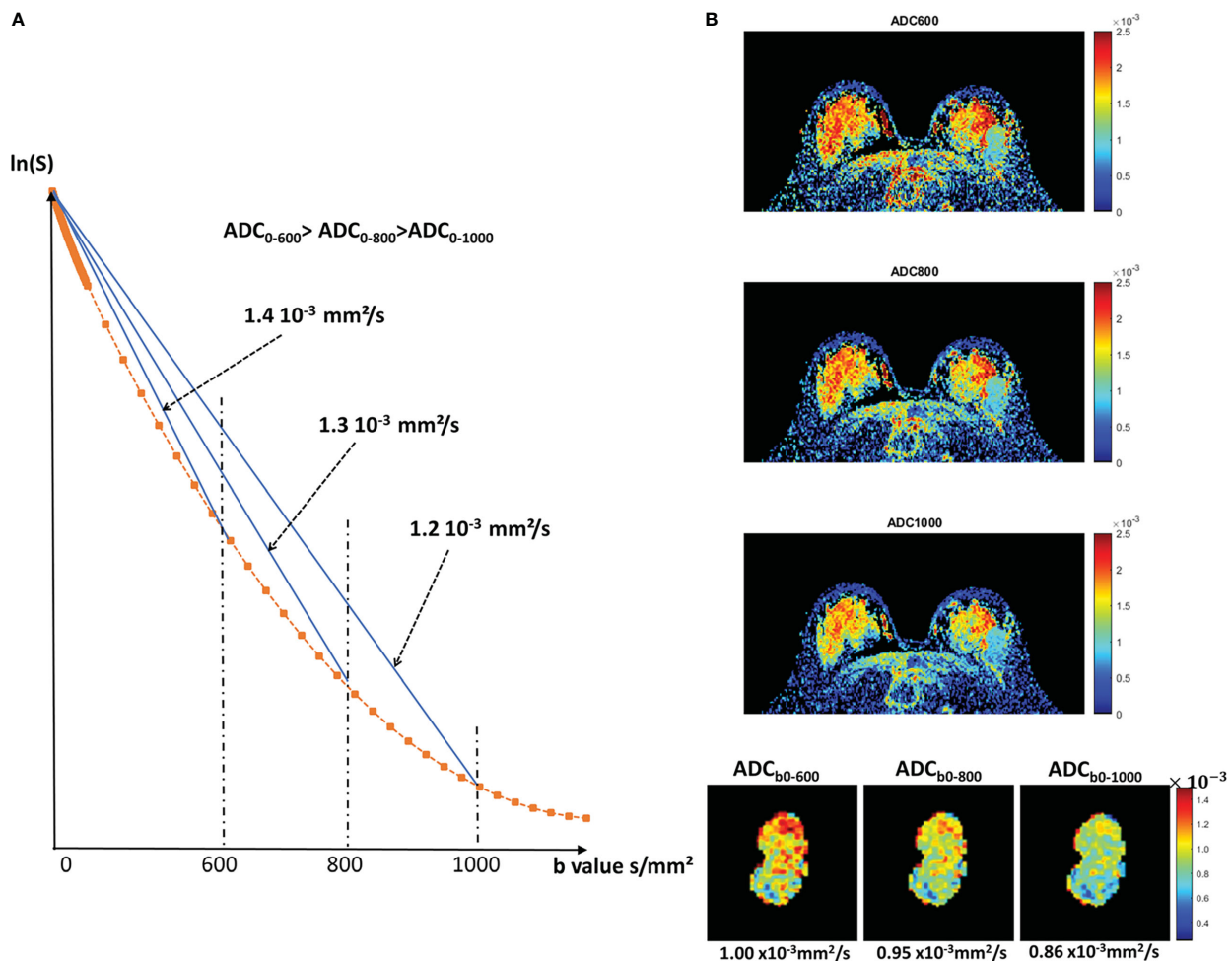


FIGURE 2

Diffusion attenuation versus b value. (A) Signal attenuation as a function of b value (logarithmic scale). With free diffusion we expect a straight line, whose slope is the diffusion coefficient. In tissues, diffusion is not free (non-Gaussian), resulting in a curvature. Therefore, the ADC taken from $b=0$ and any b values will decrease when b increases. The effects of IVIM, which result in a curvature at very low b values ($<200 \text{ s/mm}^2$), are not shown for clarity. (B) Example of a breast tumor showing that indeed the ADC value decreases when using larger b values (reprinted with permission from 32).

diffusion”) which leads to an attenuation of the signal in the presence of diffusion encoding gradient pulses. This effect has been named IntraVoxel Incoherent Motion (IVIM) (35). In the presence of blood microcirculation, the global attenuation of the MRI signal, $S(b)/S(0)$, becomes the sum of two components, one for tissue diffusion and one for the blood compartment:

$$S(b)/S_0 = f_{\text{IVIM}} \exp[-b(D^* + D_{\text{blood}})] + (1 - f_{\text{IVIM}}) \exp(-bD) \quad (2)$$

where f_{IVIM} is the fraction of circulating blood, D^* is the pseudo-diffusion coefficient attributed to the random microcirculation of blood, D is the diffusion coefficient of water in tissue, and D_{blood} is the diffusion coefficient of water in blood. The perfusion effect is observed only at low values of b, because the pseudo-diffusion coefficient, D^* , associated with blood flow is higher than the water diffusion coefficient and decreases more rapidly with the b-value.

IVIM MRI has become an important modality for perfusion imaging, with applications throughout the body (31, 36), particularly in cancer imaging (detection of neovascularization and treatment efficacy). A key feature of IVIM diffusion MRI is that it does not involve contrast agents, and it may serve as an attractive alternative to perfusion MRI in

some patients with contraindications to contrast agents, or in patients with renal insufficiency at risk for NSF (see above).

3.1.3 Non-Gaussian diffusion

Another important feature of diffusion MRI, which should be considered, counter-intuitively, as an advantage and not as a limitation, is that the ADC value depends on the acquisition parameters, especially the b-value, because diffusion in tissues is not “free” but “hindered”. With free (or “Gaussian”) diffusion, as in a cyst, the ADC remains the same regardless of the set of b values used to measure it (only the accuracy of ADC estimates changes with b values). However, in most tissues, the ADC value decreases as the diffusion sensitivity is increased by the b value (32) (Figure 2).

The reason is that an increasing number of molecules slowed down by their interaction with microstructural tissue components (fibers, cell membranes) during their diffusion movements become visible in the highly diffusion-sensitized MRI signal. This non-Gaussian diffusion behavior is therefore more pronounced when high b values are used. In short, sticking to the “optimal” b-value (e.g. 800 s/mm^2) deprives one of the potentially valuable clinical

information about tissue microstructure encoded in the “non-Gaussian diffusion” provided by higher b-values. To reveal this hidden information about tissue microstructure, one must rely on models other than the standard ADC. There are essentially two types of such models. Some approaches aim to model the diffusion MRI signal biophysically, based on the different tissue compartments present in the tissue, as with NODDI (Neurite Orientation Dispersion and Density Imaging) used in the brain (37). The other way is simply to model the decay of the scattering signal mathematically, empirically, without any assumptions about the underlying biophysical properties of the tissue. Although several models have been proposed (38), the most popular approach simply quantifies the deviation of the scattering signal behavior from an ideal Gaussian behavior. This is the so-called Kurtosis model (39), also called Diffusion Kurtosis Imaging, DKI (40). With the Kurtosis model, which also includes the IVIM effect, the signal is described as follows:

$$S(b)/S_0 = f_{IVIM} \exp[-b(D^* + D_{\text{blood}})] + (1 - f_{IVIM}) \exp[-b \text{ADCo} + (b \text{ADCo})^2 K/6] \quad (3)$$

ADCo is the extrapolated ADC value as b approaches 0 and K is the Kurtosis quantifying the deviation from Gaussian scattering ($K=0$ for Gaussian diffusion). Kurtosis has shown great potential for characterizing pathological or physiological conditions (41). A major drawback of DKI, however, is that it requires the acquisition of large data sets with multiple values of b to be fitted with equation (3), which significantly increases acquisition times, a premium in clinical practice.

3.1.4 Abbreviated quantitative diffusion MRI protocols

However, it is possible to obtain quantitative information about non-Gaussian diffusion with data sets acquired with a limited range of b values. For example, using data acquired for only 2 b values, one can calculate a shifted ADC (sADC). The concept of sADC (31) is based on the use of shifted key b-values (200 and 1500 s/mm² for the breast, instead of 0 and 800 s/mm²) providing an interesting balance between Gaussian and non-Gaussian diffusion effects. This approach has been evaluated for the breast (42). Another approach, S-index, provides a direct classification of tissue types by calculating a distance between the acquired signals and a library of reference (“signature”) signals from known or simulated tissues (e.g., benign, malignant, etc.) by intrinsically accounting for Gaussian and non-Gaussian diffusion effects, without the need for any mathematical or biophysical modeling (43). This approach has also been shown to provide the immunohistochemical status and molecular subtypes of invasive breast carcinomas (44) (Figure 3).

It is also possible to estimate the main parameters of the IVIM/Kurtosis model described by equation [3], f_{IVIM} , ADCo and K, without fitting, using a limited set of 4 b values (b_0 , b_1 , b_2 and b_3 in ascending order), providing the signal:noise ratio is not too low. The proposed algorithm assumes that IVIM effects become negligible in signals acquired above b_1 and that non-Gaussian diffusion effects appear visible in b_2 and b_3 signals. According to this 4b algorithm the model parameters estimates can be calculated as:

$$f_{IVIM} \approx 1 - \exp[-(D_1 - D_2 - H)(b_1.b_2)/(b_2 - b_1)] \quad (4)$$

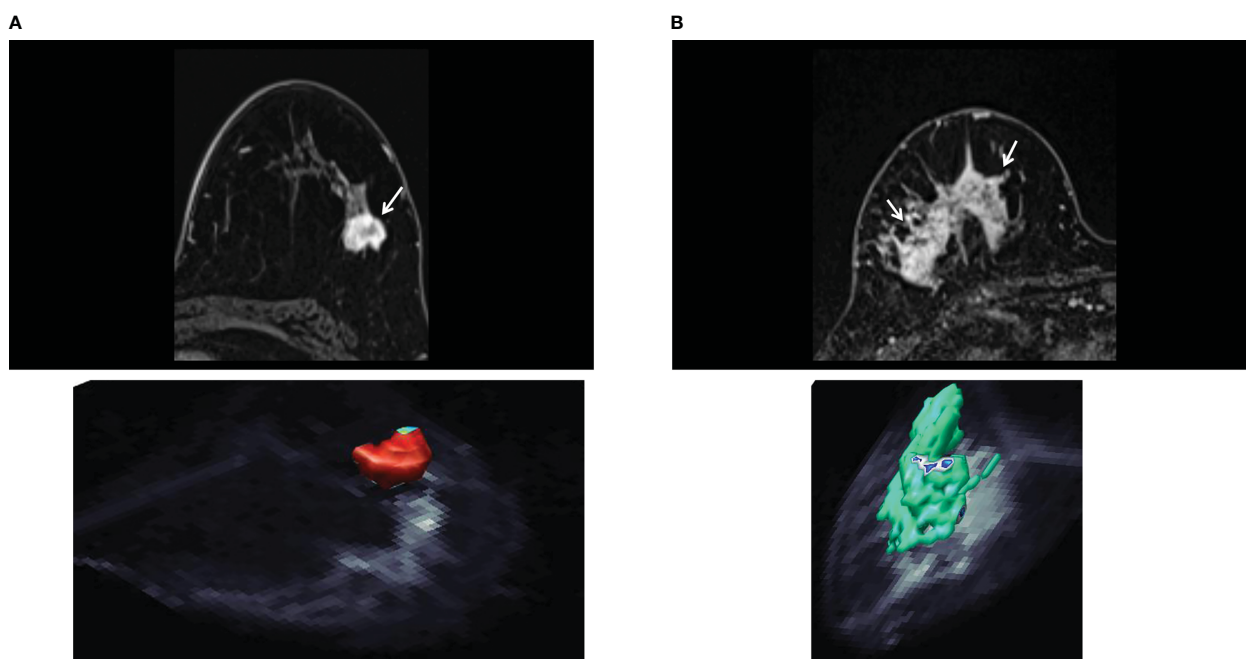


FIGURE 3
S-index. (A) Invasive ductal carcinoma of luminal A type in 50-year-old woman. The axial early-phase DCE-MRI image shows a mass with an irregular margin (top, arrow). The three-dimensional rendering voxel-by-voxel S-index image (bottom) shows the entire tumor in reddish color, corresponding to an average signature index (S-index) of this mass was 90.4. (B) Human epidermal growth factor receptor 2 enriched invasive ductal carcinoma in 73-year-old woman. The axial early-phase DCE-MRI image shows non-mass enhancement with a heterogeneous internal pattern in the right breast (top, arrow). The three-dimensional S-index rendering map of the entire tumor exhibits a yellow-green color (mean S-index of 55.8). [Adapted from (44)].

$$sADC = \ln[S(b_1)/S(b_3)] / (b_3 - b_1) \quad (5)$$

$$ADCo \approx sADC + (b_1 \cdot b_3) A \quad (6)$$

$$K \approx 6A / ADCo^2 \quad (7)$$

where $D1 = \ln[S(0)/S(b_1)] / (b_1 - b_0)$; $D2 = \ln[S(0)/S(b_2)] / (b_2 - b_0)$; $D3 = \ln[S(0)/S(b_3)] / (b_3 - b_0)$; $H = (D2 - D3) / (b_3 - b_2) + \ln(1 - F) / (b_3 \cdot b_2)$; $F = 1 - \exp[-(D1 - D2)(b_1 \cdot b_2) / (b_2 - b_1)]$; $A = (D2 - D3) / (b_3 - b_2) + \ln(1 - f_{IVIM}) / (b_3 \cdot b_2)$.

In the absence of IVIM and non-Gaussian diffusion effects one obviously has $D1 = D2 = D3 = sADC$, $ADCo = sADC$ and $K = 0$.

If non-Gaussian diffusion is present without IVIM effects ($f_{IVIM} = 0$) $ADCo$ and K are obtained exactly as:

$$ADCo(\text{no IVIM}) = sADC + (D2 - D3) (b_1 \cdot b_3) / (b_3 - b_2)$$

$$K(\text{no IVIM}) = 6(D2 - D3) / [(b_3 - b_2) ADCo^2]$$

A graphical interpretation of this set of equations can be given by plotting the (log) of the (curved) signal attenuation

versus the b value and the straight lines corresponding to $D1$, $D2$, $D3$ and $sADC$ (Figure 4). In the absence of IVIM and non-Gaussian diffusion effects the signal attenuation follows a straight line with a slope $D1 = D2 = D3 = sADC = ADCo$. In the presence of IVIM effects only the curvature at low b values creates an angle between the $D1$ and $D2$ lines. From this angle f_{IVIM} can be estimated (Eq. 4) while $ADCo$ remains very close to the $sADC$ (Eq. 6). In the presence of non-Gaussian diffusion only the curvature at high b values forms an angle between the $D2$ and $D3$ lines, from which K can be estimated (Eq. 7). However, one can see that f_{IVIM} and non-Gaussian diffusion slightly contribute also to the angle between the $D2$ and $D3$ lines, and the angle between the $D1$ and $D2$ lines, respectively. Hence, estimated f_{IVIM} and K values must be corrected (variables H and A in Eq. 4, 6). Note that the $sADC$ now includes $ADCo$, f_{IVIM} and K effects, so that the $ADCo$ values derived from $sADC$ must be corrected (A variable). Also, with this algorithm D^* cannot be estimated, however, a review of the literature shows that D^* is a parameter difficult to estimate even with the full fitting approach, resulting in extremely variable clinical relevance.

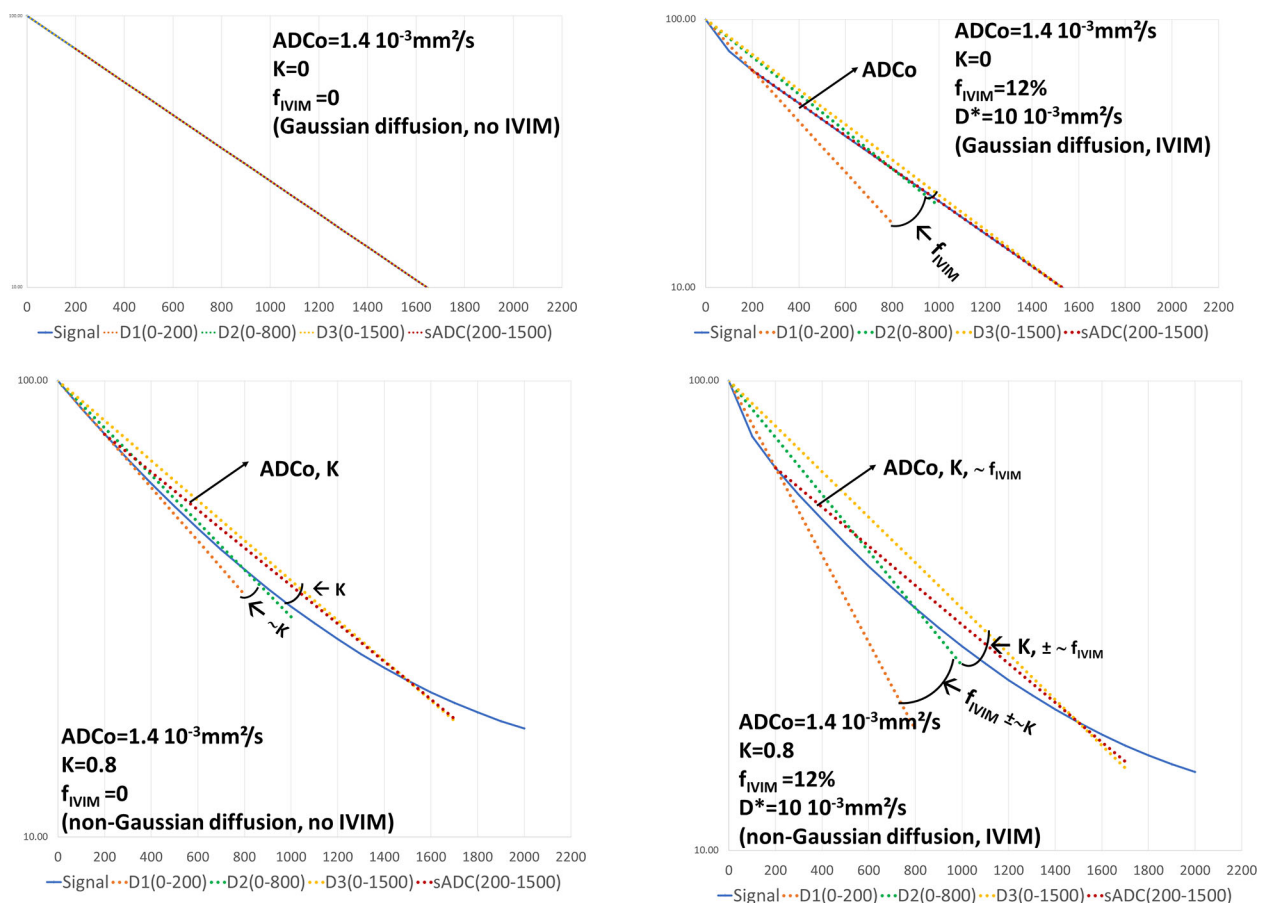


FIGURE 4

Graphical representation of the 4b-diffusion MRI abbreviated protocol. The plots show the signal attenuation and the straight lines associated to the intermediate calculation parameters (slopes) $D1$, $D2$, $D3$ used to estimate f_{IVIM} , $ADCo$ and K , as well as $sADC$. In the presence of free diffusion ($K=0$) and in the absence of IVIM effects (top left) the signal attenuation follows a straight line whose slope is $ADCo$. $D1$, $D2$, $D3$ and $sADC$ are all equal to $ADCo$. When IVIM effects appear (top right) the $D1$ line starts to deviate from the signal attenuation curve with an angle with $D2$ reflecting f_{IVIM} . With non-Gaussian diffusion effects only, both $D1$ and $D3$ deviate from $D2$ with an angle reflecting K (bottom left). The $sADC$ line depends on $ADCo$ and K . When both IVIM and non-Gaussian diffusion effects are present one can see that the angle between $D1$ and $D2$ primarily reflects f_{IVIM} while the $D2D3$ angle mainly reflects K (bottom right). The $sADC$ now reflects $ADCo$, K and f_{IVIM} . By combining $sADC$ with $D1$, $D2$ and $D3$ one can get accurate estimates of $ADCo$, K and f_{IVIM} using equation (4-7).

It is expected that those abbreviated quantitative DWI protocols will play a major role, in addition to qualitative DWI, such as DWIBS (see below), in the context of breast cancer screening with diffusion MRI.

3.1.5 Diffusion tensor imaging

Molecular mobility in biological tissues may not be the same in all directions, which is referred to as diffusion anisotropy. In the breast diffusion anisotropy can arise from the geometric organization of the glandular tissue around ducts. To characterize the effects of anisotropy, diffusion-weighted images must be sensitized to diffusion along multiple directions (at least 6) within the Diffusion Tensor Imaging (DTI) (45). With DTI one gets information on the tissue mean diffusivity, MD, which is equivalent to an orientation invariant ADC, and lambda values (λ_1 , λ_2 , and λ_3) which give diffusivity along the main diffusivity directions (so-called eigenvectors e_1 , e_2 , and e_3). The eigenvector e_1 , associated with the highest λ value, λ_1 , is aligned along the main orientation of aligned structures (e.g. ducts), allowing to produce maps showing their orientation in space. Some vendors propose to estimate the MD from a set of 3 orthogonal directions, but this is an approximation that should not be used in the presence of strong anisotropy effects. The genuine mean diffusivity is simply the average of the 3 λ values. The other important parameter, called Fractional Anisotropy (FA), quantifies the degree of anisotropy (FA = 0 indicates that diffusion is isotropic). It is calculated from the λ values. Whereas the existence of diffusion anisotropy in fibroglandular breast tissue has been claimed by many groups (46–49), the nature of the anatomical features which might cause this anisotropy remains somewhat controversial. Some studies have shown that breast cancer lesions could be associated with significantly lower FA values relative to normal breast tissue, and that λ_1 or ($\lambda_1 - \lambda_3$) could overperform the ADC (or MD) for lesion detection and classification (50, 51). However, one has to keep in mind that λ_1 and λ_3 (and FA which depends on them) are, by principle, highly sensitive to noise because of the strongly non-linear nature of the DTI calculation algorithm. The mere fact that MD values are lower in malignant lesions than in normal tissue might lead to reduced FA values, which should not necessarily be interpreted as “reduced anisotropy” (32).

3.2 Clinical performance of breast diffusion MRI

There is an extensive literature on breast diffusion MRI. We give below a brief summary of the highlights. Many more details can be found in (10). In addition, a survey of the implementation of breast DWI in clinical practice from the EUSOBI has recently been published (52).

3.2.1 Qualitative lesion detection

Most often diffusion MRI is used qualitatively for lesion detection. Lesion detection can be achieved from DWI acquired with high b values, which have a higher contrast between breast lesions (which appear bright) and normal parenchyma (dark background). Breast cancer detection using DWI has been shown to be more sensitive than

MMG, with the DWI screening approach allowing to detect mammographically occult cancers (53–55) and DWI has been shown to detect significantly more contralateral breast cancers in women with unilateral breast cancer than MMG (56). High b-values are also useful in decreasing false-positive breast cancer cases (57).

A variant of the DWI techniques for qualitative lesion detection is DWIBS (Diffusion-weighted Whole-body Imaging with Background body signal). A previous study in 280 patients has shown that the diagnostic performance using non-contrast technique including DWIBS for breast lesion detection (sensitivity, specificity, diagnostic accuracy, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) values of 94%, 79%, 86%, 79% and 94%, respectively) was comparable to that of DCE-MRI (sensitivity, specificity, diagnostic accuracy, PPV and NPV values of 98%, 83%, 90%, 84% and 98%, respectively) (58). DWIBS performed with Maximum Intensity Projection (MIP) mapping also has a comparable diagnostic performance (sensitivity, specificity, PPV and NPV values of 92%, 94%, 93%, and 92%, respectively) to that of DCE-MRI performed with MIP (sensitivity, specificity, PPV and NPV values of 85%, 90%, 89%, and 87%, respectively). MIP-DWIBS has been shown to rule out previously suggested malignancy on screening MMGs in 50 participants with carcinoma in 24 patients (59).

3.2.2 Quantitative lesion evaluation

Nevertheless, a unique feature of breast DWI is its quantitative assessment capability. As the most popular quantitative marker, the ADC can be used as a threshold to sort out benign from malignant lesions (13, 60, 61), but also to build a lexicon to describe and classify lesions, for instance to distinguish breast cancers from benign lesions (13). Many groups have also found significant differences of ADC values between benign and metastatic breast lymph nodes (62–64), however, their diagnostic performance in differentiating these lymph nodes still need further investigation (64) compared to simpler markers such as the lymph node size. Quantitative DWI in addition to DCE-MRI and other plain MRI such as T1WI and T2WI also leads to improved diagnostic performance, in terms of specificity for BI-RADS (Breast Imaging-Reporting And Data System) 3 and 4 lesions, or evaluating malignancies with BI-RADS 4 lesions (65, 66).

DWI is often used in multiparametric protocols in combination with other MRI modalities, such as DCE-MRI, contributing to improve overall diagnostic specificity and accuracy over DCE-MRI alone (67), especially when examining non-Gaussian diffusion (42) (Figure 5). The combination of DCE-MRI and DWI could increase diagnostic accuracy in characterization of non-mass-like enhancement lesions (68). It has also been reported that DWI combined with T2WI improved the diagnostic specificity of enhancing lesions incidentally detected in breast DCE-MRI (69), and that multiparametric DWI outcome parameters have associations with molecular prognostic factors or subtypes (70, 71).

Advanced diffusion markers (IVIM, Kurtosis, DTI) can further increase diagnostic performance, although there are not yet used in routine clinical practice (Figures 3, 5). IVIM parameters have been shown to provide a high diagnostic performance in differentiating benign and malignant breast tumors (sensitivity = 86%, specificity = 86%, AUC = 0.91 for D, sensitivity = 80%, specificity = 76%, AUC = 0.85 for f, and sensitivity = 84%, specificity = 59%, AUC = 0.71 for D*) (72), especially in combination with DCE-MRI (73), and IVIM

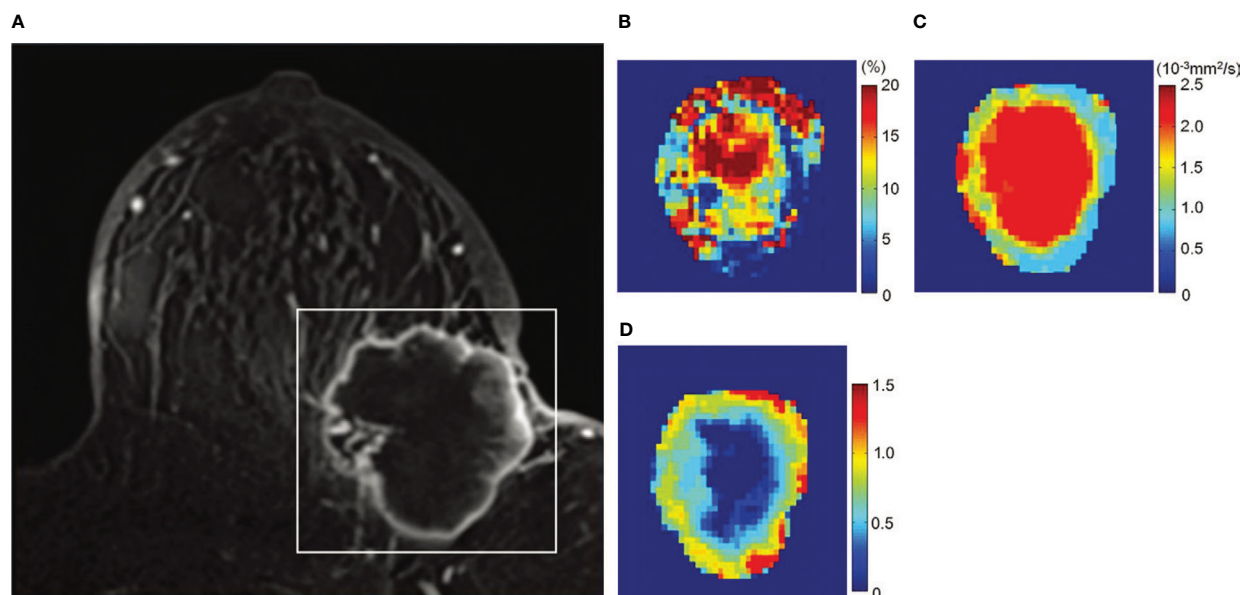


FIGURE 5

Example of non-Gaussian diffusion MRI maps in breast lesions. Images in a 72-year-old woman with invasive ductal carcinoma. (A) Dynamic contrast-enhanced axial MR image, (B) f_{IVIM} map, (C) ADC_0 map, and (D) K map. The white rectangle on (A) shows the area covered by the parametric maps. (B) f_{IVIM} distribution is heterogeneous inside the tumor. The peripheral area of the tumor exhibits low ADC_0 (C) and high K values (D), suggesting high cellularity diffusion hindrance effect (likely from cellular membranes) corresponding to the viable malignant component (also high f_{IVIM} values), whereas the central part shows high ADC_0 and low K, suggesting lower cellularity (possible necrosis with fluid motion at the center resulting in high f_{IVIM} values). [Adapted from (42)].

parameters are known to be correlated with DCE-MRI parameters (74). IVIM histogram parameters have been shown to be associated with molecular prognostic factors (75, 76). Regarding DKI higher K and lower MD values are usually observed in malignant compared to benign lesions (41, 77), DKI was found to be useful in the differentiation of additional suspicious lesions at preoperative breast MRI (78). In 2 recent meta-analyses of DKI studies (79, 80) the sensitivity and specificity of K and MD to differentiate malignant from benign breast lesions were found to be around 89-90% and 86-88% for K, and 84-86% and 83-88% for MD. The utility of DKI in differentiating molecular prognostic factors (81) or predicting treatment response (82) has also been reported.

As for DTI parameters, malignant breast lesions have significantly lower MD and λ_1 values compared to benign lesions (46). Indeed, λ_1 and MD are known to have a high diagnostic performance in differentiating malignant and benign breast lesions (AUC 0.97, sensitivity 93%, specificity 92% for λ_1 and AUC 0.92, sensitivity 87%, specificity 83% for D (50), although this trend might not be related to diffusion anisotropy [see above (32)]. Similarly, the use of FA to differentiate between malignant and benign breast tumors remains controversial, suggesting that caution should be exercised in the use of this parameter, although several studies have suggested its usefulness to sort out malignant and benign lesions (14, 50). Yamaguchi et al. (83) also reported higher FA in lesions with more favorable prognostic factors, such as positive estrogen receptor status, lower nuclear grade and cancer intrinsic subtype, and reduced DTI metrics had association with poor prognostic factors of breast cancer (84). A correlation between DTI parameters and molecular prognostic factors (estrogen receptor status or Ki-67) has been shown (49) and DTI has been investigated to differentiate recurrent

breast cancer from post operative changes with breast-conserving surgery in patients (85).

3.2.3 Diffusion MRI as a stand-alone modality

In reviewing the literature, the overall sensitivity of DWI alone is very high, approaching 90% for detection of breast malignancy (with a specificity of approximately 82%, superior to any other imaging modality, including GBCA MRI) (51).

Many investigators have studied the potential of DWI alone for non-contrast cancer detection. The performance of DWI imaging for cancer detection is variable across studies, with a mean sensitivity of 81% (range 44-97%) and a mean specificity of 88% (range 73-96%) (86). This variation could be due to the diversity of the study population as well as the image acquisition protocols, highlighting the need for standardization (see below). Nonetheless, DWI based primarily on qualitative assessment is less sensitive than DCE MRI (mean sensitivity of 80 vs. 90s % for DWI vs. DCE MRI in studies (54, 58, 59, 67, 87-93). This situation is entertaining the idea that diffusion MRI would be difficult as a stand-alone modality compared to DCE-MRI and multiparametric MRI (67). However, in reviewing the literature, the overall sensitivity of quantitative ADC alone is very high, approaching 90% for detection of breast malignancy (with a specificity of approximately 82%, superior to any other imaging modality, including GBCA-MRI) (51). In any case, with respect to screening, diffusion MRI offers much better clinical performance than MMG or ultrasound (94). This is an important point, considering that repeated use of GBCA would be a problem for screening. Not only does DWI remain more sensitive than MMG across studies (52, 92, 95, 96), but mammography-occluded breast cancers are better represented with DWI than with ultrasound (94).

4 How breast diffusion MRI could be implemented to give optimal performance in breast cancer screening

4.1 Standardization

Despite this good clinical performance, it may seem surprising that DWI has not yet been recommended to be used as a stand-alone modality for breast cancer evaluation, let alone for breast cancer screening. Diffusion MRI is not even included in the BI-RADS lexicon used to assess breast lesions from GBCA MRI (97), although it is considered useful (52). The main reason is likely the high variability of the results found in the literature, especially with regard to ADC values (13, 98). The EUSOBI international committee on breast DWI has provided guidelines for obtaining optimized and consistent results (13). This report, along with the EUSOBI survey (52) have pointed out to an urgent need for standardization of DWI acquisition and processing protocols to achieve consistent results among breast DWI users.

Technical advances in MRI scanners, particularly for gradient hardware and fast imaging, facilitate the exploration of new features beyond ADC by allowing perfusion-driven IVIM to become more reliable (99), providing access to non-Gaussian diffusion through high b-values, and investigating diffusion time effects. This increasing flexibility of diffusion MRI acquisitions is supporting the expansion of more complex models, allowing for a better understanding of the relationship between diffusion MRI parameters and the microscopic characteristics of the underlying tissue. This is particularly true in the field of breast imaging, where a wide variety of diffusion MRI techniques have great potential for clinical applications in the breast field. However, this flexibility implies that some normalization must be implemented in order to compare quantitative results obtained at multiple sites. Not only are ADC values strongly dependent on b-values (100–102), but they are also influenced by TE, due to differences in T2 values between tissue components. Hidden parameters, such as diffusion time (set by the duration and intervals of the gradient pulses) also have important effects. For example, while high performance gradient hardware can achieve high b-values with shorter TEs, increasing the signal-to-noise ratio, diffusion contrast may be partially lost, as diffusion hindrance decreases with short diffusion time (103). Thus, there is a clear need for standardization of acquisition protocols. Validation of these protocols in different clinical sites would benefit from calibrated phantoms, as suggested by EUSOBI (the European Society of Breast Imaging) (13), QIBA (Quantitative Imaging Biomarkers Alliance) organized by the Radiological Society of North America (RSNA) (104). Clearly, additional efforts are needed in collaboration with vendors if consensus is to be reached on optimal acquisition parameters for diffusion MRI of the breast (10).

4.2 Technical requirements and improvements

4.2.1 Image acquisition

Single-shot echo-planar imaging (EPI) is currently the method of choice for *in vivo* diffusion imaging, as it allows efficient and ultrafast

acquisition of multiple diffusion-weighted images (different b-values) without in-plane motion artifacts, to which diffusion MRI is notoriously sensitive. Nevertheless, EPI has several limitations related to spatial resolution, artifacts, and signal-to-noise ratio. In particular, small breast lesions (<2 mm) may be undetectable. In addition, EPI requires a very homogeneous magnetic field. For breast imaging, field inhomogeneities may be more pronounced at the air/tissue interface in the anterior part of the breast, resulting in local image distortion or signal loss. Another source of geometric distortion comes from eddy currents induced by the switching of strong diffusion encoding gradient pulses. Therefore, the degree of geometric distortion increases with the b-value. This geometric distortion must be corrected before performing any quantitative analysis involving multiple values of b to avoid artifacts around small lesions, especially at high spatial resolution. Segmented EPI acquisitions (e.g., ‘RESOLVE’ (Readout Segmentation of Long Variable Echo-trains) (105) can overcome these limitations at the cost of longer diffusion times and a sensitivity to motion between acquired segments that must be corrected using *ad-hoc* approaches during image reconstruction. Parallel acquisition techniques, which allow simultaneous signal collection using an array of multiple RF coils, can also address these limitations. Incorrect fat suppression can also lead to misinterpretation of diffusion MRI, as residual fat present in breast tumors results in low diffusion values, mimicking malignancy, visually and quantitatively (ADC values). The Spectrally Adiabatic Inversion Recovery (SPAIR) method has been recommended for breast imaging (13).

4.2.2 Image processing

Efforts are also needed on the image processing side. Diffusion-weighted images are often noisy, especially for high b-values, because the signal is strongly attenuated by the diffusion effect. Noise is a vicious enemy because it is not always visible, while having a profound impact on the values of the parameters estimated with the various models available including ADC. For high b-values, due to the nature of the MRI signal (a “magnitude” signal that cannot be negative), there always remains a background noise signal and the diffusion signal remains above a threshold, the “noise floor”, instead of asymptotically approaching 0, resulting in underestimated ADC values. If one classifies lesions (e.g., benign or malignant) on the basis of ADC threshold values, it is easy to see that this trap of underestimated ADC could lead to a significant bias toward the “malignant” nature of lesions. Therefore, an adequate signal-to-noise ratio must be ensured, e.g., by increasing the voxel size (at the expense of spatial resolution) or by repeating image acquisitions at high b-values for signal averaging before amplitude reconstruction (which unfortunately increases acquisition time). Finally, background noise effects must also be removed from the signals before analysis, especially in images acquired at high b-values (106). Noise effects may partly explain the discrepancies in the literature on the different reported values of diffusion MRI and IVIM parameters. Image preprocessing could also include steps to correct for motion artifacts and geometric distortion before the signals can be processed to calculate ADC values or estimate parameters for advanced DWI models. Another problem with clinical diffusion MRI is that quantitative analysis is often performed remotely on workstations and not on the acquisition console, which is

cumbersome. Efforts are underway by vendors to provide dedicated tools for breast DWI [see the final chapters of the book (10)].

DWI data analysis would also benefit from recent developments in artificial intelligence (AI). Various approaches are being investigated for breast MRI, as well as remarkably increasing applications of convolutional neural network models (107) and machine learning (108). For example, a recent study showed that DWI radiomic classifiers for differentiating suspicious lesions in 50 asymptomatic women screened with MMG outperformed the average ADC, with an area under the curve (AUC) of receiver operating characteristics (ROC) of 84.2%/85.1% for unconstrained/constrained radiomic classifiers compared with 77.4% for the average ADC (109). The AI-based multiparametric MRI approach, including DCE, T2WI, and DWI, had better diagnostic performance (AUC ROC area of 0.852) than ultrafast DCE alone (0.811) (110). Machine learning with multiparametric MRI (DCE, DWI, and T2WI) also found that several features, including those of DWI (minimum ADC), were relevant features for predicting residual cancer burden (111). Whole breast segmentation on DWI data from different institutions and scanner types was also found to be effective using deep learning methods, which could facilitate computer-assisted quantitative analyses of DWI images of the breast (112).

4.3 Toward a low-cost, dedicated MRI system for breast cancer screening

Given the outstanding clinical performance of breast MRI, which has much higher sensitivity and specificity than MMG and does not rely on x-rays, it should ideally be the screening modality of choice for many women. Unfortunately, breast MRI remains expensive when performed using general-purpose body MRI scanners operating at 1.5T or even 3T. The cost (and limited availability) of these scanners prohibits the use of MRI as a screening modality (the cost today is approximately \$1000 for a 40-minute exam). Breast cancer screening with MRI is therefore reserved for women with moderate to high risk of breast cancer, as detailed above. However, if a small-scale, inexpensive, dedicated breast MRI scanner were available, MRI could be used as a screening imaging modality, rather than MMG, for more women, such as women with dense breasts or a family history of breast cancer.

One issue that comes to mind when considering breast cancer screening with MRI is the use of GBCA, as examinations will need to be repeated over many years, knowing that an accumulation of gadolinium deposits in the brain or other organs in patients who have received multiple injections of contrast agents has been demonstrated. For this reason, several groups have considered the possibility of using diffusion MRI as a stand-alone imaging modality for breast cancer screening (55, 59). As detailed above, diffusion MRI, which is completely noninvasive, has been successfully used to differentiate between benign and malignant breast lesions and tumor extension. Diffusion MRI also has the potential to detect many occult mammographic and clinical carcinomas of the breast, making it a preferred modality for cancer screening. Contrast agents could then still be used, but as a second line if necessary.

A major technical implication of using diffusion MRI instead of GBCA MRI is that only one breast can be scanned at a time, as with MMG, making the design of a dedicated breast MRI scanner much easier, smaller, and therefore available at much lower cost. Here we propose some specifications that might be kept in mind when designing such a dedicated imaging system. Ideally, the device should be small to be mobile and affordable. In total, the footprint of the system should also be small compared to the 5-gauss line. Patients could be in a standing position, as a bed structure would increase space and cost (Figure 6). This will also shorten the examination time and therefore reduce imaging costs. The disadvantage is that breast motion (which is already a problem with conventional MRI) will have to be controlled mechanically (motion sensors) and/or using *ad hoc* post-processing algorithms. Field homogeneity should be < 1ppm/20cm peak-to-peak (0.05 ppm after shimming). This is a very important requirement because breast MRI requires “fat suppression” techniques that rely on the differential frequencies between fat and water resonance frequencies. In addition, thoracic bones and air contained in the lungs are responsible for local magnetic susceptibility effects that distort the magnetic field. As with general MRI, the field stability must be better than 0.05ppm/h (10^{-4} ppm/10 minutes). An open design will also allow image-guided biopsy or therapy (113). The field strength should be low to keep construction and maintenance costs as low as possible, ideally using helium-free magnets. This means that several technical improvements must be implemented to maintain sufficient signal-to-noise ratios, especially when using high diffusion weighting (large b values). Efficient and powerful gradient hardware must be implemented to achieve high b-values while maintaining a short TE. Innovative radio frequency systems will need to be designed for both transmission and reception. For example, receive coil arrays could be tailored to different breast sizes to maximize fill factor, such as “bra coils.” AI algorithms that have been developed for acquisition (sparse sampling) and signal processing (114) will help maintain adequate signal-to-noise levels while achieving spatial resolution greater than 2 mm. To exploit the full content of the diffusion MRI signal, one can even envision that processing will be performed not on the reconstructed images (which are only for the eyes of radiologists and clinicians), but on the denoised raw signals using AI algorithms trained and optimized to detect disease signatures. The images will then be reconstructed by focusing on these anomalies when they are detected. Assuming that no suspicious lesions will be found in the vast majority of cases, radiologists will be able to focus on the remaining cases that the AI system will identify as difficult to classify.

Clearly, designing such a prototype is a team effort. Clinicians must work closely with physicists, engineers and technicians, not only to design the most patient-friendly system, but also with market attention. The price of the overall system should be similar to that of high-end mammography systems, around 400 k€. In addition to the cost of building a proof-of-concept prototype, costs for patenting, multi-center trials and market research, calibration and quality control, FDA (US Food and Drug Administration) and CE (Conformité Européenne) marking, etc., must be considered. We sincerely hope that some vendors will be interested in this challenge, invest and bring such a breast MRI screening device to the market for the benefit of patients worldwide.

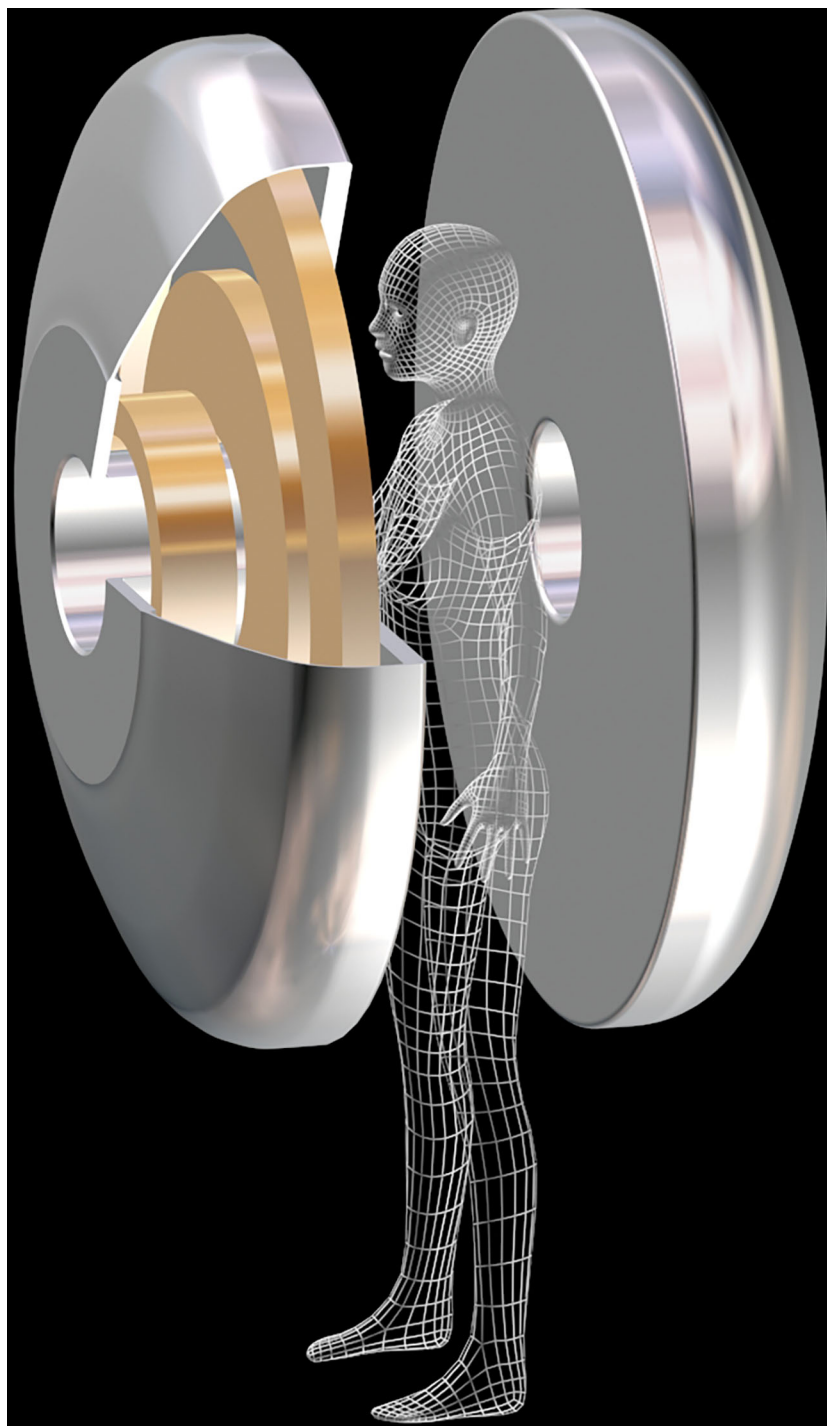


FIGURE 6
Prototype of a MRI magnet dedicated to breast cancer screening. The superconducting magnet consists in 2 halves. The patient stands between the 2 halves. To reduce size and cost, this magnet can be tailored for scanning one breast at a time (courtesy T. Schild, Ifre/CEA).

5 Conclusion

Non-contrast breast diffusion MRI has emerged as a potential alternative for breast cancer screening and lesion characterization. Without GBCA injections and with higher sensitivity and

specificity than MMG, breast diffusion MRI is emerging as an ideal imaging modality for cancer screening. Consensus is needed to define the population categories that could benefit from this approach, such as women at moderate to high risk for cancer. Efforts are still needed to standardize acquisition and processing

protocols and to decrease the cost of breast MRI examinations. To this end, the development of a low-cost MRI system dedicated to DWI for breast cancer screening is an option that should be seriously considered.

Author contributions

Conception or design of the work. MI, DB. Materials collection. MI, DB. Materials analysis and interpretation. MI, DB. Drafting the article. MI, DB. Critical revision of the article. MI, DB. Final approval of the version to be published. MI, DB. All authors contributed to the article and approved the submitted version.

Funding

This study was supported in-part by the Program for the Development of Next-generation Leading Scientists with Global Insight (L-INSIGHT), sponsored by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* (2022) 72:7–33. doi: 10.3322/caac.21708
2. Casasent AK, Edgerton M, Navin NE. Genome evolution in ductal carcinoma in situ: Invasion of the clones. *J Pathol* (2017) 241:208–18. doi: 10.1002/path.4840
3. van Luijt PA, Heijnsdijk EAM, Fracheboud J, Overbeek LIH, Broeders MJM, Wesseling J, et al. The distribution of ductal carcinoma *in situ* (DCIS) grade in 4232 women and its impact on overdiagnosis in breast cancer screening. *Breast Cancer Res* (2016) 18:47. doi: 10.1186/s13058-016-0705-5
4. Park HL, Chang J, Lal G, Lal K, Ziogas A, Anton-Culver H. Trends in treatment patterns and clinical outcomes in young women diagnosed with ductal carcinoma in situ. *Clin Breast Cancer* (2018) 18:e179–85. doi: 10.1016/j.clbc.2017.08.001
5. Kuhl CK, Schrading S, Leutner CC, Morakkabati-Spitz N, Wardelmann E, Fimmers R, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol* (2005) 23:8469–76. doi: 10.1200/JCO.2004.00.4960
6. Pinder SE, Shaaban A, Deb R, Desai A, Gandhi A, Lee AHS, et al. NHS Breast screening multidisciplinary working group guidelines for the diagnosis and management of breast lesions of uncertain malignant potential on core biopsy (B3 lesions). *Clin Radiol* (2018) 73:682–92. doi: 10.1016/j.crad.2018.04.004
7. Giannotti E, James JJ, Chen Y, Sun R, Karupiah A, Yemm J, et al. Effectiveness of percutaneous vacuum-assisted excision (VAE) of breast lesions of uncertain malignant potential (B3 lesions) as an alternative to open surgical biopsy. *Eur Radiol* (2021) 31:9540–7. doi: 10.1007/s00330-021-08060-z
8. Bianchi S, Caimi S, Renne G, Cassano E, Ambrogetti D, Cattani MG, et al. Positive predictive value for malignancy on surgical excision of breast lesions of uncertain malignant potential (B3) diagnosed by stereotactic vacuum-assisted needle core biopsy (VANCb): A large multi-institutional study in Italy. *Breast* (2011) 20:264–70. doi: 10.1016/j.breast.2010.12.003
9. Lehman CD. Magnetic resonance imaging in the evaluation of ductal carcinoma in situ. *J Natl Cancer Inst Monogr* (2010) 2010:150–1. doi: 10.1093/jncimonographs/lgq030
10. Iima M, Partridge S, Le Bihan D. *Diffusion MRI of the Breast*. Philadelphia: Elsevier Health Sciences. Philadelphia (2022). p. 260.
11. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med* (2019) 381:2091–102. doi: 10.1056/NEJMoa1903986
12. Sardaneli F, Schiaffino S, Cozzi A, Carbonaro LA. “Gadolinium-based contrast agents for breast MRI and uncertainties about brain gadolinium retention.”. *Breast MRI High-risk Screening*. (2020) p:63–82. doi: 10.1007/978-3-030-41207-4_5
13. Baltzer P, Mann RM, Iima M, Sigmund EE, Clauser P, Gilbert FJ, et al. Diffusion-weighted imaging of the breast—a consensus and mission statement from the EUSOBI international breast diffusion-weighted imaging working group. *Eur Radiol* (2020) 30:1436–50. doi: 10.1007/s00330-019-06510-3
14. Lotter W, Diab AR, Haslam B, Kim JG, Grisot G, Wu E, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat Med* (2021) 27:244–9. doi: 10.1038/s41591-020-11174-9
15. Lee MV, Aharon S, Kim K, Sunn Konstantinoff K, Appleton CM, Stwalley D, et al. Recent trends in screening breast MRI. *J Breast Imaging* (2022) 4:39–47. doi: 10.1093/jbi/wbab088
16. Kaiser W. MRI Of the female breast. *First Clin results. Arch Int Physiol Biochim* (1985) 93:67–76.
17. Knopp MV, Weiss E, Sinn HP, Mattern J, Junkermann H, Radeleff J, et al. Pathophysiologic basis of contrast enhancement in breast tumors. *J Magn Reson Imaging* (1999) 10:260–6. doi: 10.1002/(SICI)1522-2586(199909)10:3<260::AID-JMRI6>3.0.CO;2-7
18. Nam SY, Ko ES, Lim Y, Han B-K, Ko EY, Choi JS, et al. Preoperative dynamic breast magnetic resonance imaging kinetic features using computer-aided diagnosis: Association with survival outcome and tumor aggressiveness in patients with invasive breast cancer. *PloS One* (2018) 13:e0195756. doi: 10.1371/journal.pone.0195756
19. Kuhl CK, Mielcareck P, Klaschik S, Leutner C, Wardelmann E, Gieseke J, et al. Dynamic breast MR imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology* (1999) 211:101–10. doi: 10.1148/radiology.211.1.r99ap38101
20. Leithner D, Wengert GJ, Helbich TH, Thakur S, Ochoa-Albiztegui RE, Morris EA, et al. Clinical role of breast MRI now and going forward. *Clin Radiol* (2018) 73:700–14. doi: 10.1016/j.crad.2017.10.021
21. Clauser P, Mann R, Athanasiou A, Prosch H, Pinker K, Dietzel M, et al. A survey by the European society of breast imaging on the utilisation of breast MRI in clinical practice. *Eur Radiol* (2018) 28:1909–18. doi: 10.1007/s00330-017-5121-4
22. Weinstein SP, Slanetz PJ, Lewin AA, Battaglia T, Chagpar AB, Dayaratna S, et al. ACR appropriateness criteria® supplemental breast cancer screening based on breast density. *J Am Coll Radiol* (2021) 18:S456–73. doi: 10.1016/j.jacr.2021.09.002
23. Rahman WT, Helvie MA. Breast cancer screening in average and high-risk women. *Best Pract Res Clin Obstet Gynaecol* (2022) 83:3–14. doi: 10.1016/j.bpobgyn.2021.11.007
24. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers R-D, Bieling HB. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection—a novel approach to breast cancer screening with MRI. *J Clin Oncol* (2014) 32:2304–10. doi: 10.1200/JCO.2013.52.5386
25. High WA, Ayers RA, Cowper SE. Gadolinium is quantifiable within the tissue of patients with nephrogenic systemic fibrosis. *J Am Acad Dermatol* (2007) 56:710–2. doi: 10.1016/j.jaad.2007.01.022
26. ACR Manual on Contrast Media. (2013). Available at: http://www.acr.org/media/ACR/Documents/PDF/QualitySafety/Resources/Contrast%20Manual/2013_Contrast_Media.pdf.
27. Attari H, Cao Y, Elmholt TR, Zhao Y, Prince MR. A systematic review of 639 patients with biopsy-confirmed nephrogenic systemic fibrosis. *Radiology* (2019) 292:376–86. doi: 10.1148/radiol.2019182916

Acknowledgments

We would like to thank Ms. Rena Nakayama for contributing to edit this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

28. Kanda T, Ishii K, Kawaguchi H, Kitajima K, Takenaka D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. *Radiology* (2014) 270:834–41. doi: 10.1148/radiol.13131669
29. Sardaneli F, Cozzi A, Trimboli RM, Schiaffino S. Gadolinium retention and breast MRI screening: More harm than good? *Am J Roentgenology* (2020) 214:324–7. doi: 10.2214/AJR.19.21988
30. Melsaether AN, Kim E, Mema E, Babb J, Kim SG. Preliminary study: Breast cancers can be well seen on 3T breast MRI with a half-dose of gadobutrol. *Clin Imaging* (2019) 58:84–9. doi: 10.1016/j.clinimag.2019.06.014
31. Iima M, Le Bihan D. Clinical intravoxel incoherent motion and diffusion MRI imaging: Past, present, and future. *Radiology* (2016) 278:13–32. doi: 10.1148/radiol.2015150244
32. Iima M, Partridge SC, Le Bihan D. Six DWI questions you always wanted to know but were afraid to ask: Clinical relevance for breast diffusion MRI. *Eur Radiol* (2020) 30:2561–70. doi: 10.1007/s00330-019-06648-0
33. Le Bihan D. Apparent diffusion coefficient and beyond: What diffusion MR imaging can tell us about tissue structure. *Radiology* (2013) 268:318–22. doi: 10.1148/radiol.13130420
34. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. *Radiology* (1986) 161:401–7. doi: 10.1148/radiology.161.2.3763909
35. Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* (1988) 168:497–505. doi: 10.1148/radiology.168.2.3393671
36. Le Bihan D, Iima M, Federau C. *Sigmund EE eds. (2018) intravoxel incoherent motion (IVIM) MRI: Pan Stanford publishing*. Singapore. p. 534.
37. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* (2012) 61:1000–16. doi: 10.1016/j.neuroimage.2012.03.072
38. Tang L, Zhou XJ. Diffusion MRI of cancer: From low to high b-values. *J Magn Reson Imaging* (2019) 49:23–40. doi: 10.1002/jmri.26293
39. Chabert S, Meca C, Le Bihan D. Relevance of the information about the diffusion distribution in vivo given by kurtosis in q-space imaging. *Proc 12th Annu Meeting ISMRM* (2004) 1238.
40. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn Reson Med* (2005) 53:1432–40. doi: 10.1002/mrm.20508
41. Wu D, Li G, Zhang J, Chang S, Hu J, Dai Y. Characterization of breast tumors using diffusion kurtosis imaging (DKI). *PLoS One* (2014) 9:e113240. doi: 10.1371/journal.pone.0113240
42. Iima M, Kataoka M, Kanao S, Onishi N, Kawai M, Ohashi A, et al. Intravoxel incoherent motion and quantitative non-Gaussian diffusion MR imaging: Evaluation of the diagnostic and prognostic value of several markers of malignant and benign breast lesions. *Radiology* (2018) 287:432–41. doi: 10.1148/radiol.2017162853
43. Goto M, Le Bihan D, Yoshida M, Sakai K, Yamada K. Adding a model-free diffusion MRI marker to BI-RADS assessment improves specificity for diagnosing breast lesions. *Radiology* (2019) 292:84–93. doi: 10.1148/radiol.2019181780
44. Goto M, Le Bihan D, Sakai K, Yamada K. The diffusion MRI signature index is highly correlated with immunohistochemical status and molecular subtype of invasive breast carcinoma. *Eur Radiol* (2022) 32:4879–88. doi: 10.1007/s00330-022-08562-4
45. Furman-Haran E, Grobeld D, Nissan N, Shapiro-Feinberg M, Degani H. Can diffusion tensor anisotropy indices assist in breast cancer detection? *J Magn Reson Imaging* (2016) 44:1624–32. doi: 10.1002/jmri.25292
46. Luo J, Hippe DS, Rahbar H, Parsian S, Rendi MH, Partridge SC. Diffusion tensor imaging for characterizing tumor microstructure and improving diagnostic performance on breast MRI: A prospective observational study. *Breast Cancer Res* (2019) 21:102. doi: 10.1186/s13058-019-1183-3
47. Eyal E, Shapiro-Feinberg M, Furman-Haran E, Grobeld D, Golan T, Itzhak Y, et al. Parametric diffusion tensor imaging of the breast. *Invest Radiol* (2012) 47:284–91. doi: 10.1097/RLI.0b013e3182438e5d
48. Baltzer PAT, Schäfer A, Dietzel M, Grässel D, Gajda M, Camara O, et al. Diffusion tensor magnetic resonance imaging of the breast: A pilot study. *Eur Radiol* (2011) 21:1–10. doi: 10.1007/s00330-010-1901-9
49. Onaygil C, Kaya H, Ugurlu MU, Aribal E. Diagnostic performance of diffusion tensor imaging parameters in breast cancer and correlation with the prognostic factors. *J Magn Reson Imaging* (2017) 45:660–72. doi: 10.1002/jmri.25481
50. Wang K, Li Z, Wu Z, Zheng Y, Zeng S, Linning E, et al. Diagnostic performance of diffusion tensor imaging for characterizing breast tumors: A comprehensive meta-analysis. *Front Oncol* (2019) 9:1229. doi: 10.3389/fonc.2019.01229
51. Baxter GC, Graves MJ, Gilbert FJ, Patterson AJ. A meta-analysis of the diagnostic performance of diffusion MRI for breast lesion characterization. *Radiology* (2019) 291:632–41. doi: 10.1148/radiol.2019182510
52. Lo Gullo R, Sevilmedu V, Baltzer P, Le Bihan D, Camps-Herrero J, Clauser P, et al. A survey by the European society of breast imaging on the implementation of breast diffusion-weighted imaging in clinical practice. *Eur Radiol* (2022) 32:6588–97. doi: 10.1007/s00330-022-08833-0
53. Partridge SC, Demartini WB, Kurland BF, Eby PR, White SW, Lehman CD. Differential diagnosis of mammographically and clinically occult breast lesions on diffusion-weighted MRI. *J Magn Reson Imaging* (2010) 31:562–70. doi: 10.1002/jmri.22078
54. Yabuuchi H, Matsuo Y, Sunami S, Kamitani T, Kawanami S, Setoguchi T, et al. Detection of non-palpable breast cancer in asymptomatic women by using unenhanced diffusion-weighted and T2-weighted MR imaging: Comparison with mammography and dynamic contrast-enhanced MR imaging. *Eur Radiol* (2011) 21:11–7. doi: 10.1007/s00330-010-1890-8
55. McDonald ES, Hammersley JA, Chou SHS, Rahbar H, Scheel JR, Lee CI, et al. “Performance of DWI as a rapid unenhanced technique for detecting mammographically occult breast cancer in elevated-risk women with dense breasts.” *AJR Am J Roentgenol* (2016) 207(1):205–16. doi: 10.2214/AJR.15.15873
56. Ha SM, Chang JM, Lee SH, Kim ES, Kim SY, Kim YS, et al. Detection of contralateral breast cancer using diffusion-weighted magnetic resonance imaging in women with newly diagnosed breast cancer: Comparison with combined mammography and whole-breast ultrasound. *Korean J Radiol* (2021) 22:867–79. doi: 10.3348/kjr.2020.1183
57. Ochi M, Kuroiwa T, Sunami S, Murakami J, Miyahara S, Nagaie T, et al. Diffusion-weighted imaging (b value = 1500 s/mm²) is useful to decrease false-positive breast cancer cases due to fibrocystic changes. *Breast Cancer* (2013) 20:137–44. doi: 10.1007/s12282-011-0319-9
58. Telegrafo M, Rella L, Stabile Ianora AA, Angelelli G, Moschetta M. Unenhanced breast MRI (STIR, T2-weighted TSE, DWIBS): An accurate and alternative strategy for detecting and differentiating breast lesions. *Magn Reson Imaging* (2015) 33:951–5. doi: 10.1016/j.mri.2015.06.002
59. Bickelhaupt S, Laun FB, Tesdorff J, Lederer W, Daniel H, Stieber A, et al. Fast and noninvasive characterization of suspicious lesions detected at breast cancer X-ray screening: Capability of diffusion-weighted MR imaging with MIPs. *Radiology* (2016) 278:689–97. doi: 10.1148/radiol.2015150425
60. Iima M, Honda M, Sigmund EE, Ohno Kishimoto A, Kataoka M, Togashi K, et al. Diffusion MRI of the breast: Current status and future directions. *J Magn Reson Imaging* (2020) 52:70–90.
61. He M, Ruan H, Ma M, Zhang Z. Application of diffusion weighted imaging techniques for differentiating benign and malignant breast lesions. *Front Oncol* (2021) 11:694634. doi: 10.3389/fonc.2021.694634
62. Fardanesh R, Thakur SB, Sevilmedu V, Horvat JV, Gullo RL, Reiner JS, et al. Differentiation between benign and metastatic breast lymph nodes using apparent diffusion coefficients. *Front Oncol* (2022) 12:795265. doi: 10.3389/fonc.2022.795265
63. De Cataldo C, Bruno F, Palumbo P, Di Sibio A, Arrigoni F, Clemente A, et al. Apparent diffusion coefficient magnetic resonance imaging (ADC-MRI) in the axillary breast cancer lymph node metastasis detection: A narrative review. *Gland Surg* (2020) 9:2225–34. doi: 10.21037/gs-20-546
64. Iima M, Kataoka M, Okumura R, Togashi K. Detection of axillary lymph node metastasis with diffusion-weighted MR imaging. *Clin Imaging* (2014) 38:633–6. doi: 10.1016/j.clinimag.2014.04.016
65. Dijkstra H, Dorrius MD, Wielema M, Pijnappel RM, Oudkerk M, Sijens PE. Quantitative DWI implemented after DCE-MRI yields increased specificity for BI-RADS 3 and 4 breast lesions. *J Magn Reson Imaging* (2016) 44:1642–9. doi: 10.1002/jmri.25331
66. Sun SY, Ding Y, Li Z, Nie L, Liao C, Liu Y, et al. Multiparameter MRI model with DCE-MRI, DWI, and synthetic MRI improves the diagnostic performance of BI-RADS 4 lesions. *Front Oncol* (2021) 11:699127. doi: 10.3389/fonc.2021.699127
67. Pinker K, Moy L, Sutton EJ, Mann RM, Weber M, Thakur SB, et al. Diffusion-weighted imaging with apparent diffusion coefficient mapping for breast cancer detection as a stand-alone parameter: Comparison with dynamic contrast-enhanced and multiparametric magnetic resonance imaging. *Invest Radiol* (2018) 53:587–95. doi: 10.1097/RLI.0000000000000465
68. Yabuuchi H, Matsuo Y, Kamitani T, Setoguchi T, Okafuji T, Soeda H, et al. Non-mass-like enhancement on contrast-enhanced breast MR imaging: Lesion characterization using combination of dynamic contrast-enhanced and diffusion-weighted MR images. *Eur J Radiol* (2010) 75:e126–32. doi: 10.1016/j.ejrad.2009.09.013
69. Arponen O, Masarwah A, Sutela A, Taina M, Könönen M, Sironen R, et al. Incidentally detected enhancing lesions found in breast MRI: Analysis of apparent diffusion coefficient and T2 signal intensity significantly improves specificity. *Eur Radiol* (2016) 26:4361–70. doi: 10.1007/s00330-016-4326-2
70. Suo S, Cheng F, Cao M, Kang J, Wang M, Hua J, et al. Multiparametric diffusion-weighted imaging in breast lesions: Association with pathologic diagnosis and prognostic factors. *J Magn Reson Imaging* (2017) 46:740–50. doi: 10.1002/jmri.25612
71. Sharma U, Sah RG, Agarwal K, Parshad R, Seenu V, Mathur SR, et al. Potential of diffusion-weighted imaging in the characterization of malignant, benign, and healthy breast tissues and molecular subtypes of breast cancer. *Front Oncol* (2016) 6:126. doi: 10.3389/fonc.2016.00126
72. Liang J, Zeng S, Li Z, Kong Y, Meng T, Zhou C, et al. Intravoxel incoherent motion diffusion-weighted imaging for quantitative differentiation of breast tumors: A meta-analysis. *Front Oncol* (2020) 10:585486. doi: 10.3389/fonc.2020.585486
73. Wang Q, Guo Y, Zhang J, Wang Z, Huang M, Zhang Y. Contribution of IVIM to conventional dynamic contrast-enhanced and diffusion-weighted MRI in differentiating benign from malignant breast masses. *Breast Care* (2016) 11:254–8. doi: 10.1159/000447765
74. Liu C, Wang K, Chan Q, Liu Z, Zhang J, He H, et al. Intravoxel incoherent motion MR imaging for breast lesions: comparison and correlation with pharmacokinetic evaluation from dynamic contrast-enhanced MR imaging. *Eur Radiol* (2016) 26:3888–98. doi: 10.1007/s00330-016-4241-6
75. Cho GY, Moy L, Kim SG, Baete SH, Moccaldi M, Babb JS, et al. Evaluation of breast cancer using intravoxel incoherent motion (IVIM) histogram analysis: Comparison with

- malignant status, histological subtype, and molecular prognostic factors. *Eur Radiol* (2016) 26:2547–58. doi: 10.1007/s00330-015-4087-3
76. Lee YJ, Kim SH, Kang BJ, Kang YJ, Yoo H, Yoo J, et al. Intravoxel incoherent motion (IVIM)-derived parameters in diffusion-weighted MRI: Associations with prognostic factors in invasive ductal carcinoma. *J Magn Reson Imaging* (2017) 45:1394–406. doi: 10.1002/jmri.25514
77. Sun K, Chen X, Chai W, Fei X, Fu C, Yan X, et al. Breast cancer: Diffusion kurtosis MR imaging-diagnostic accuracy and correlation with clinical-pathologic factors. *Radiology* (2015) 277:46–55. doi: 10.1148/radiol.15141625
78. Park VY, Kim SG, Kim E-K, Moon HJ, Yoon JH, Kim MJ. Diffusional kurtosis imaging for differentiation of additional suspicious lesions on preoperative breast MRI of patients with known breast cancer. *Magn Reson Imaging* (2019) 62:199–208. doi: 10.1016/j.mri.2019.07.011
79. Li Z, Li X, Peng C, Dai W, Huang H, Li X, et al. The diagnostic performance of diffusion kurtosis imaging in the characterization of breast tumors: A meta-analysis. *Front Oncol* (2020) 10:575272. doi: 10.3389/fonc.2020.575272
80. Gu H, Cui W, Luo S, Deng X. Diagnostic performance of diffusion kurtosis imaging for benign and malignant breast lesions: A systematic review and meta-analysis. *Appl Bionics Biomech* (2022) 2022:2042736. doi: 10.1155/2022/2042736
81. Yang ZL, Li Y, Zhan CA, Hu YQ, Guo YH, Xia LM, et al. Evaluation of suspicious breast lesions with diffusion kurtosis MR imaging and connection with prognostic factors. *Eur J Radiol* (2021) 145:110014. doi: 10.1016/j.ejrad.2021.110014
82. Zhang D, Geng X, Suo S, Zhuang Z, Gu Y, Hua J. The predictive value of DKI in breast cancer: Does tumour subtype affect pathological response evaluations? *Magn Reson Imaging* (2022) 85:28–34. doi: 10.1016/j.mri.2021.10.013
83. Yamaguchi K, Nakazono T, Egashira R, Komori Y, Nakamura J, Noguchi T, et al. Diagnostic performance of diffusion tensor imaging with readout-segmented echo-planar imaging for invasive breast cancer: Correlation of ADC and FA with pathological prognostic markers. *Magn Reson Med Sci* (2017) 16:245–52. doi: 10.2463/mrms.mp.2016-0037
84. Kim JY, Kim JJ, Kim S, Choo KS, Kim A, Kang T, et al. Diffusion tensor magnetic resonance imaging of breast cancer: Associations between diffusion metrics and histological prognostic factors. *Eur Radiol* (2018) 28:3185–93. doi: 10.1007/s00330-018-5429-8
85. Abdel Razek AAK, Zaky M, Bayoumi D, Taman S, Abdelwahab K, Alghandour R. Diffusion tensor imaging parameters in differentiation recurrent breast cancer from post-operative changes in patients with breast-conserving surgery. *Eur J Radiol* (2019) 111:76–80. doi: 10.1016/j.ejrad.2018.12.022
86. Shin HJ, Moon WK, Amornsiripanitch N, Partridge SC. DIFFUSION MRI OF THE BREAST. In: Iima M, Partridge S, Le Bihan D, editors. *Diffusion MRI as a stand-alone unenhanced approach for breast imaging and screening*. Philadelphia: Elsevier Health Sciences (2022). p. 260.
87. Baltzer PAT, Benndorf M, Dietzel M, Gajda M, Runnebaum IB, Kaiser WA. False-positive findings at contrast-enhanced breast MRI: A BI-RADS descriptor study. *AJR Am J Roentgenol* (2010) 194(6):1658–63. doi: 10.2214/AJR.09.3486
88. Zhang L, Tang M, Min Z, Lu J, Lei X, Zhang X. Accuracy of combined dynamic contrast-enhanced magnetic resonance imaging and diffusion-weighted imaging for breast cancer detection: A meta-analysis. *Acta Radiol* (2016) 57:651–60. doi: 10.1177/0284185115597265
89. O'Flynn EAM, Blackledge M, Collins D, Downey K, Doran S, Patel H, et al. Evaluating the diagnostic sensitivity of computed diffusion-weighted MR imaging in the detection of breast cancer. *J Magn Reson Imaging* (2016) 44:130–7. doi: 10.1002/jmri.25131
90. Kang JW, Shin HJ, Shin KC, Chae EY, Choi WJ, Cha JH, et al. Unenhanced magnetic resonance screening using fused diffusion-weighted imaging and maximum-intensity projection in patients with a personal history of breast cancer: Role of fused DWI for postoperative screening. *Breast Cancer Res Treat* (2017) 165:119–28. doi: 10.1007/s10549-017-4322-5
91. Baltzer PAT, Bickel H, Spick C, Wengert G, Woitek R, Kapetas P, et al. Potential of noncontrast magnetic resonance imaging with diffusion-weighted imaging in characterization of breast lesions. *Invest Radiol* (2018) 53:229–35. doi: 10.1097/RLI.0000000000000433
92. Bu Y, Xia J, Joseph B, Zhao X, Xu M, Yu Y, et al. Non-contrast MRI for breast screening: preliminary study on detectability of benign and malignant lesions in women with dense breasts. *Breast Cancer Res Treat* (2019) 177:629–39. doi: 10.1007/s10549-019-05342-5
93. Ha SM, Chang JM, Lee SH, Kim ES, Kim S-Y, Cho N, et al. Diffusion-weighted MRI at 3.0 T for detection of occult disease in the contralateral breast in women with newly diagnosed breast cancer. *Breast Cancer Res Treat* (2020) 182:283–97. doi: 10.1007/s10549-020-05697-0
94. Amornsiripanitch N, Rahbar H, Kitsch AE, Lam DL, Weitzel B, Partridge SC. Visibility of mammographically occult breast cancer on diffusion-weighted MRI versus ultrasound. *Clin Imaging* (2018) 49:37–43. doi: 10.1016/j.clinimag.2017.10.017
95. Yoshikawa MI, Ohsumi S, Sugata S, Kataoka M, Takashima S, Kikuchi K, et al. Comparison of breast cancer detection by diffusion-weighted magnetic resonance imaging and mammography. *Radiat Med* (2007) 25:218–23. doi: 10.1007/s11604-007-0128-4
96. Kazama T, Kuroki Y, Kikuchi M, Sato Y, Nagashima T, Miyazawa Y, et al. Diffusion-weighted MRI as an adjunct to mammography in women under 50 years of age: an initial study. *J Magn Reson Imaging* (2012) 36:139–44. doi: 10.1002/jmri.23626
97. Breast imaging reporting and data system. In: *ACR BI-RADS® atlas 5th edition*. American College of Radiology.
98. Newitt DC, Zhang Z, Gibbs JE, Partridge SC, Chenevert TL, Rosen MA, et al. Test-retest repeatability and reproducibility of ADC measures by breast DWI: Results from the ACRIN 6698 trial. *J Magn Reson Imaging* (2019) 49:1617–28. doi: 10.1002/jmri.26539
99. Chen W, Zhang J, Long D, Wang Z, Zhu JM. Optimization of intra-voxel incoherent motion measurement in diffusion-weighted imaging of breast cancer. *J Appl Clin Med Phys* (2017) 18:191–9. doi: 10.1002/acm2.12065
100. Pereira FPA, Martins G, Figueiredo E, Domingues MNA, Domingues RC, da Fonseca LMB, et al. Assessment of breast lesions with diffusion-weighted MRI: Comparing the use of different b values. *AJR Am J Roentgenol* (2009) 193:1030–5. doi: 10.2214/AJR.09.2522
101. Peters NHGM, Vincken KL, van den Bosch MAAJ, Luijten PR, Mali WPTM, Bartels LW. Quantitative diffusion weighted imaging for differentiation of benign and malignant breast lesions: The influence of the choice of b-values. *J Magn Reson Imaging* (2010) 31:1100–5. doi: 10.1002/jmri.22152
102. Eghtedari M, Ma J, Fox P, Guvenc I, Yang WT, Dogan BE. Effects of magnetic field strength and b value on the sensitivity and specificity of quantitative breast diffusion-weighted MRI. *Quant Imaging Med Surg* (2016) 6:374–80. doi: 10.21037/qims.2016.07.06
103. Iima M, Kataoka M, Honda M, Ohashi A, Ohno Kishimoto A, Ota R, et al. Rate of apparent diffusion coefficient change with diffusion time on breast diffusion-weighted imaging depends on breast tumor types and molecular prognostic biomarker expression. *Invest Radiol* (2021) 56:501–8. doi: 10.1097/RLI.0000000000000766
104. QIBA Profile. *Diffusion-weighted magnetic resonance imaging (DWI)*. Available at: http://qibawiki.rsna.org/images/1/1d/QIBADWIProfilev1.45_20170427_v5_accepted.pdf.
105. Porter DA, Heidemann RM. High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition. *Magn Reson Med* (2009) 62:468–75. doi: 10.1002/mrm.22024
106. Iima M, Yano K, Kataoka M, Umehana M, Murata K, Kanao S, et al. Quantitative non-Gaussian diffusion and intravoxel incoherent motion magnetic resonance imaging: differentiation of malignant and benign breast lesions. *Invest Radiol* (2015) 50:205–11. doi: 10.1097/RLI.0000000000000094
107. Yin H, Jiang Y, Xu Z, Huang W, Chen T, Lin G. Apparent diffusion coefficient-based convolutional neural network model can be better than sole diffusion-weighted magnetic resonance imaging to improve the differentiation of invasive breast cancer from breast ductal carcinoma in situ. *Front Oncol* (2021) 11:805911. doi: 10.3389/fonc.2021.805911
108. Reig B, Heacock L, Geras KJ, Moy L. Machine learning in breast MRI. *J Magn Reson Imaging* (2020) 52:998–1018. doi: 10.1002/jmri.26852
109. Bickelhaupt S, Paech D, Kickingereder P, Steudle F, Lederer W, Daniel H, et al. Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography. *J Magn Reson Imaging* (2017) 46:604–16. doi: 10.1002/jmri.25606
110. Dalmis MU, Gubern-Mérida A, Vreemann S, Bult P, Karssemeijer N, Mann R, et al. Artificial intelligence-based classification of breast lesions imaged with a multiparametric breast MRI protocol with ultrafast DCE-MRI, T2, and DWI. *Invest Radiol* (2019) 54:325. doi: 10.1097/RLI.0000000000000544
111. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, et al. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Invest Radiol* (2019) 54:110–7. doi: 10.1097/RLI.0000000000000518
112. Zhang L, Mohamed AA, Chai R, Guo Y, Zheng B, Wu S. Automated deep learning method for whole-breast segmentation in diffusion-weighted breast MRI. *J Magn Reson Imaging* (2020) 51:635–43. doi: 10.1002/jmri.26860
113. Sattarov A, McIntyre P, Motowidlo L. High-field open MRI for breast cancer screening. *IEEE Trans Appl Supercond* (2015) 25:1–5. doi: 10.1109/TASC.2014.2377049
114. Lazarus C, Weiss P, Chaffert N, Mauconduit F, El Gueddari L, Destrieux C, et al. SPARKLING: variable-density k-space filling curves for accelerated T2*-weighted MRI. *Magn Reson Med* (2019) 81:3643–61. doi: 10.1002/mrm.27678



OPEN ACCESS

EDITED BY
Alla Reznik,
Lakehead University, Canada

REVIEWED BY
Piergiorgio Cerello,
National Institute of Nuclear Physics of
Turin, Italy
Shouliang Qi,
Northeastern University, China

*CORRESPONDENCE
Xiaoping Yin
✉ yinxiaoping78@sina.com

[†]These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 09 January 2023
ACCEPTED 16 March 2023
PUBLISHED 28 March 2023

CITATION
Cai J, Guo L, Zhu L, Xia L, Qian L,
Lure YMF and Yin X (2023) Impact
of localized fine tuning in the performance
of segmentation and classification of
lung nodules from computed tomography
scans using deep learning.
Front. Oncol. 13:1140635.
doi: 10.3389/fonc.2023.1140635

COPYRIGHT
© 2023 Cai, Guo, Zhu, Xia, Qian, Lure and
Yin. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Impact of localized fine tuning in the performance of segmentation and classification of lung nodules from computed tomography scans using deep learning

Jingwei Cai^{1,2†}, Lin Guo^{3†}, Litong Zhu^{4†}, Li Xia³, Lingjun Qian³,
Yuan-Ming Fleming Lure³ and Xiaoping Yin^{1*}

¹Radiology Department, Affiliated Hospital of Hebei University, Baoding, Hebei, China, ²Clinical Medical College, Hebei University, Baoding, Hebei, China, ³Shenzhen Zhiying Medical Imaging, Shenzhen, Guangdong, China, ⁴Department of Medicine, Queen Mary Hospital, University of Hong Kong, Hong Kong, Hong Kong SAR, China

Background: Algorithm malfunction may occur when there is a performance mismatch between the dataset with which it was developed and the dataset on which it was deployed.

Methods: A baseline segmentation algorithm and a baseline classification algorithm were developed using public dataset of Lung Image Database Consortium to detect benign and malignant nodules, and two additional external datasets (i.e., HB and XZ) including 542 cases and 486 cases were involved for the independent validation of these two algorithms. To explore the impact of localized fine tuning on the individual segmentation and classification process, the baseline algorithms were fine tuned with CT scans of HB and XZ datasets, respectively, and the performance of the fine tuned algorithms was tested to compare with the baseline algorithms.

Results: The proposed baseline algorithms of both segmentation and classification experienced a drop when directly deployed in external HB and XZ datasets. Comparing with the baseline validation results in nodule segmentation, the fine tuned segmentation algorithm obtained better performance in Dice coefficient, Intersection over Union, and Average Surface Distance in HB dataset (0.593 vs. 0.444; 0.450 vs. 0.348; 0.283 vs. 0.304) and XZ dataset (0.601 vs. 0.486; 0.482 vs. 0.378; 0.225 vs. 0.358). Similarly, comparing with the baseline validation results in benign and malignant nodule classification, the fine tuned classification algorithm had improved area under the receiver operating characteristic curve value, accuracy, and F1 score in HB dataset (0.851 vs. 0.812; 0.813 vs. 0.769; 0.852 vs. 0.822) and XZ dataset (0.724 vs. 0.668; 0.696 vs. 0.617; 0.737 vs. 0.668).

Conclusions: The external validation performance of localized fine tuned algorithms outperformed the baseline algorithms in both segmentation process and classification process, which showed that localized fine tuning may be an effective way to enable a baseline algorithm generalize to site-specific use.

KEYWORDS

segmentation, classification, lung nodules, localized fine tuning, site-specific use

1 Introduction

Lung cancer is one of the most common cancers in the world (1), which has no obvious clinical symptoms in the early stage, but is hardly cured after the onset of disease. Therefore, early diagnosis and differentiation of benign and malignant pulmonary nodules has great significance for the long-term survival of patients (2). As one of the most important means to screen lung cancer for high-risk groups (3), low-dose CT scans have been widely used in health examinations, and a large amount of CT data has created heavy workload for radiologists. Deep learning (DL) is considered as a powerful tool that have gained great achievements in the detection of benign and malignant pulmonary nodules in chest CT images (4, 5). However, in most cases, decreased performance is observed when the proposed algorithm is applied in the external tests, even with adopted and balanced validation datasets (6–9).

It has been a public concern that algorithm malfunction occurs when it is applied on external dataset that is inherently different from the training set. It may halt the possible implementation of the general model into routine clinical care if it does not have a consistent accuracy for site-specific use. To obtain a comparable external test performance to the internal tests, reported studies involving training datasets from multicenter to develop the detection algorithm demonstrated that it can either underperform (10–12) or have a comparable performance to the internal test (11, 13) without any unanimous conclusion reached, which may be explained by the differences of the datasets scale and the numbers of dataset origins (14). Using local images for model training seems to be another way to obtain a site-specific used tool for diagnosis. However, a large amount of training images is needed to develop a DL algorithm, which is challenging for those regions with lower prevalence of lung nodules, especially malignant nodules. Therefore, developing a baseline algorithm using only public dataset and then recalibrating it with local images may be an effective way to reduce site-specific bias.

It has been proved that recalibration strategy with local data is able to correct for the anticipated drop in model performance. Various studies related to recalibration method were reported, but in most cases, they are statistical prediction models focusing on updating regression coefficients, or adding new covariates for the model (15–18). To the best of our knowledge, few studies have been conducted with recalibration strategy of localized fine tuning on

imaging to separately explore its impact on the segmentation and classification process.

In the study, we conducted localized fine tuning for the baseline DL algorithm of segmentation and classification to segment and classify benign and malignant nodules. The baseline algorithms were first developed using public dataset of Lung Image Database Consortium (LIDC) (19) and then 50% of the public data was replaced with local dataset to develop the fine tuned algorithms. The performance of the fine tuned algorithms and baseline algorithms were tested and compared in multicenter datasets.

2 Methods

2.1 Patient cohorts

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) of the Affiliated Hospital of Hebei University. The informed consent from human participants was waived because this is a retrospective study, and the waiver was indicated in the IRB approval document. Three datasets were involved in the study, including a public dataset of LIDC and two collected datasets named HB and XZ, respectively. All identifications of the patient were removed.

LIDC has a total of 1018 cases (the number of patients was unknown) with annotation process performed by four radiologists. Each radiologist independently reviewed the CT images and marked lesions that belonged to one of three categories (“nodule ≥ 3 mm”, “nodule < 3 mm” and “non-nodule ≥ 3 mm”). The nodules are finally marked with 5 malignancy levels, from 1 to 5 (17). As the detection algorithm was developed for the nodule-level classification, the inclusion criteria for nodules are as follows: (1) Nodule diameter > 3 mm; (2) Nodules with score greater than 3 were included with malignant label, and nodules with score less than 3 were included with benign label; (3) Nodules with borderline median malignancy (rating = 3) were excluded; (4) Nodules with only one score were excluded. Finally, 582 cases comprising of 430 malignant nodules and 671 benign nodules were included, and they were randomly divided into training and testing set at a ratio of 8:2; the training set contained 344 malignant nodules and 536 benign nodules, and the testing set contained 86 malignant nodules and 135 benign nodules.

A total of 541 patients in HB dataset were retrospectively collected from January 2017 to June 2020, and 261 patients in XZ dataset were collected from July 2019 to May 2020. The inclusion criteria for these two datasets were: (1) The patients had typical imaging signs and pathological results of the lesions; (2) There was no surgery in the lung; (3) There was no history of malignant tumor in other part of the lung. Finally, a total of 963 nodules of HB dataset were included, comprising of 537 malignant nodules and 426 benign nodules, and a total of 785 nodules in XZ dataset with 387 malignant nodules and 398 benign nodules were also involved.

2.2 CT acquisition and image preprocessing

CT scans in HB dataset were performed using Siemens 64-row 128-slice helical CT scan and 40-row 64-slice helical CT scan (SOMATOM Definition AS, tube voltage: 100 kV, tube current: 100 mA, pitch: 1.3, slice thickness: 5.0 mm, field of view (FOV): 430 mm). CT scans in XZ dataset were performed using PHILIPS Brilliance 64-row CT scan (collimator width: 0.75mm, pitch factor: 0.1-2.0, slice thickness: 0.75-2.0mm, scanning parameters 80-140KV, 80-320mAS, A scan matrix: 512×512). All CT images were independently reviewed by two radiologists (more than 5-10 years of experience in reading CT images) using Labellng software with the annotation reference (17). If two Dice coefficient values were all greater than or at least equal to 0.95, they would be averaged as the ground truth of the image. Otherwise, a senior radiologist (more than 20 years of experience in reading CT images) would review and outline the images again to make the final determination. Since the CT images were generated by different scanning devices with different resolutions, all data were spatially resampled with the isotropic interval of $1.0 \text{ mm} \times 1.0 \text{ mm} \times 1.0 \text{ mm}$ (20).

2.3 Development of the baseline segmentation algorithm

As shown in Figure 1, 3D MaskRCNN was used to develop a baseline segmentation algorithm to detect and segment nodules on LIDC database. Before inputting the CT images to the network, the input images were transformed to physical millimeter size from pixel size with the spatial location information unchanged. The lung area was first extracted with the remaining part supplemented with pixel of 170 whose neighborhoods are close to one another, where this significantly reduces noise while preserving most image content, and then the images were randomly cropped to the size of $[128, 128, 128]$ as input training. 3D MaskRCNN is similar to the 2D MaskRCNN which consists of backbone architecture, RPN head and ROI head. The backbone architecture used in the research is resnet50, for which kernels with the size of $3 \times 3 \times 3$ were used to convolve the input image, and the feature maps output from it were input into the pooling layer to aggregate contiguous values to one scalar by the mean. The RPN architecture includes a convolutional layer and two following heads which were used to generate every anchor's shift and the score belonging to foreground, respectively. The ROI align head was involved to pool different proposals to boxes with the shape of $7 \times 7 \times 7$, and then a box header and a mask predictor were applied to finetune box position and format the lesion boundary. Specifically, in the research, the image was first input into the backbone and it would output 256 features at a $1/32$ ratio of the raw image size, then these features maps were input into the RPN network and 1000 proposals sorted by scores were obtained. Finally, the 1000 proposals were reshaped to $7 \times 7 \times 7$ boxes and all the boxes were input into mask head. We ended up selecting the predicted result with a threshold of 0.5. The total training epoch was 200, and ROI Head and Mask Head were added when the epoch was 65 and 80, respectively.

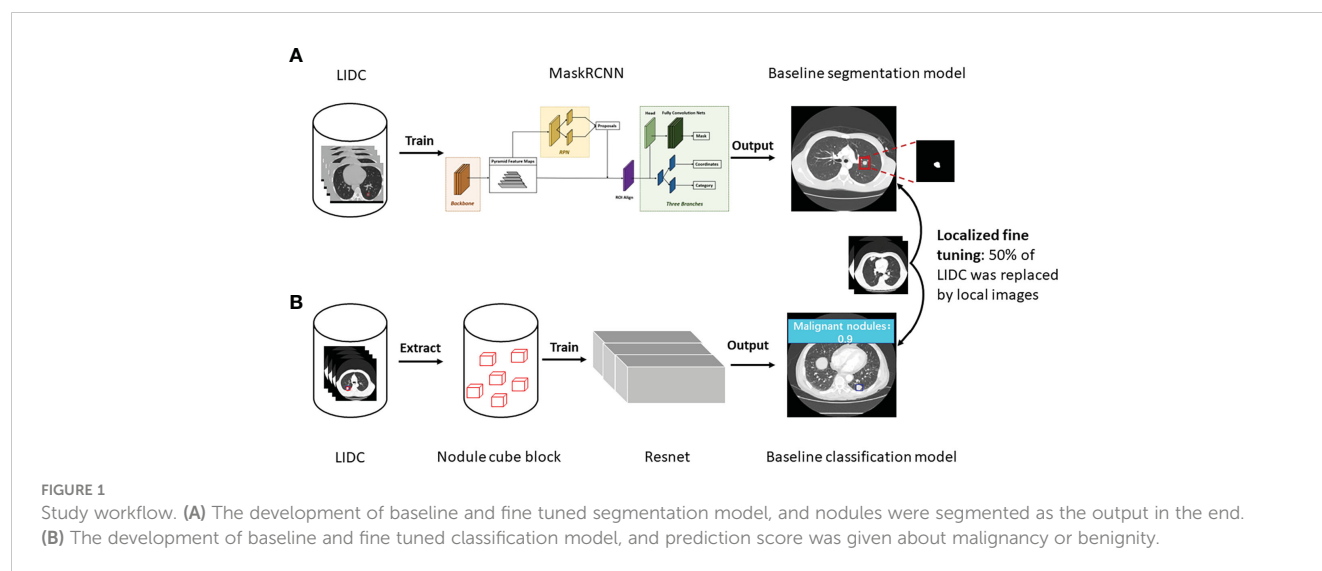


FIGURE 1

Study workflow. (A) The development of baseline and fine tuned segmentation model, and nodules were segmented as the output in the end. (B) The development of baseline and fine tuned classification model, and prediction score was given about malignancy or benignity.

2.4 Development of the baseline classification algorithm

Resnet was used to develop the baseline classification algorithm for benign and malignant nodules diagnosis (Figure 1). Specifically, first, in the binary classification task of benign and malignant nodules, the center point of the nodule was used as the reference point to extent 64 pixels in the x and y directions, and 32 layers were expanded in the z direction, forming a nodular cube block with the size of [3, 32, 64, 64], which was the input of the algorithm. Then the resnet18-3D was applied to make the calculation of the input of [b, 3, 32, 64, 64], and output [b, 2], where b is the batch size of the algorithm input.

2.5 Algorithm fine tuning

For both baseline segmentation algorithm and classification algorithm, 50% of the LIDC training set was replaced by HB and XZ datasets, and then they were trained again to be locally fine tuned, before which the HB and XZ datasets were divided into two parts of sets respectively. For the HB dataset, the one consisting of 172 malignant nodules and 268 benign nodules was used for algorithm fine tuning, and the other set consisting of 365 malignant nodules and 158 benign nodules was used as an independent test. Similarly, for XZ dataset, one set consisting of 172 malignant nodules and 268 benign nodules was used for algorithm fine tuning, and the other set consisting of 215 malignant nodules and 130 benign nodules was used as an independent test. Both baseline algorithms and fine-tuned algorithms were evaluated on HB and XZ independent sets respectively, and their performance were compared in the end (i.e., baseline segmentation algorithm vs. fine-tuned segmentation algorithm; baseline classification algorithm vs. fine-tuned

classification algorithm).

2.6 Statistical analysis

In the process of evaluating the segmentation algorithm performance, labeled nodules by radiologists are defined as positive findings, and we illustrated segmentation test results by Dice coefficient (DICE), Intersection over Union (IOU), and Average Surface Distance (ASD). For the classification results, the positive findings are malignant nodules and benign nodules are negative, and the receiver operating characteristic (ROC) curve, the value of the area under the ROC curve (AUC), accuracy, sensitivity, specificity and F1 score were used. Statistical analysis was performed using Python 3.8 and SPSS 20. Statistical tests were conducted with p -value < 0.05 as an indicator of statistical significance.

3 Results

3.1 Clinical characteristics

The main characteristics of patients in the HB and XZ datasets are shown in Figure 2. 541 patients from HB dataset were 54.2% males, and the median age was 62 years with an age range of 17–85 years. XZ included 241 patients with 50.2% males (median age of 61 years; age range 21–87 years). There was no significant difference in the patient age ($P = 0.668$) and gender ($P = 0.292$) for both cohorts. However, we observed that the distribution of benign and malignant nodules was statistically significant among LIDC, HB and XZ datasets ($P < 0.001$), and the two-two pairwise comparison between any two cohorts also showed significant difference (i.e., LIDC vs. HB: $P < 0.001$; LIDC vs. XZ: $P < 0.001$; HB vs. HB: $P = 0.007$).

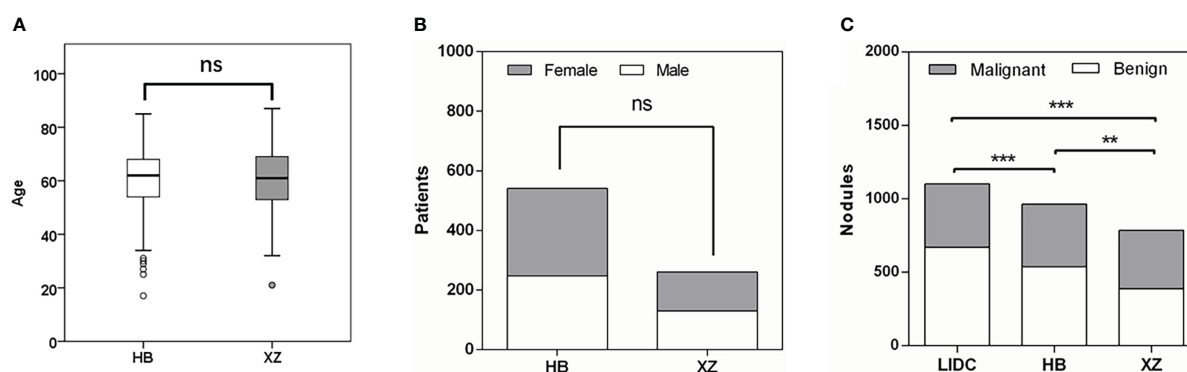


FIGURE 2 Patients characteristics. (A) Age distribution in HB and XZ datasets. (B) Gender composition in HB and XZ datasets. (C) The composition of malignant nodules and benign nodules in LIDC, HB and XZ datasets. LIDC, lung image database consortium. ns, not significant; ** p -value < 0.01; *** p -value < 0.001.

TABLE 1 Performance of baseline and fine tuned segmentation model in public LIDC dataset and two independent collected datasets.

Measure	Performance		
	Datasets		
	LIDC	HB	XZ
Dice coefficient (DICE)			
Baseline algorithm	0.771	0.444	0.486
Fine tuned algorithm	NA	0.593	0.601
Delta in DICE	NA	33.56%	23.66%
<i>P</i>	NA	0.021	0.048
Intersection over Union (IOU)			
Baseline algorithm	0.642	0.348	0.378
Fine tuned algorithm	NA	0.450	0.482
Delta in IOU	NA	29.31%	27.51%
<i>P</i>	NA	0.029	0.035
Average Surface Distance (ASD)			
Baseline algorithm	0.244	0.304	0.358
Fine tuned algorithm	NA	0.283	0.225
Delta in ASD	NA	-6.91%	-37.15%
<i>P</i>	NA	0.067	0.022

LIDC, lung image database consortium; NA, not applicable.

3.2 Effect of fine tuning on segmentation algorithms

The performance of the baseline and fine tuned segmentation algorithms assessed by the DICE, IOU, and ASD are summarized in Table 1. In the internal set of LIDC, the DICE, IOU, ASD of the baseline algorithm were 0.771, 0.642, 0.244, respectively. Then we observed a drop in its performance for external tests, with three metrics being 0.444, 0.348 and 0.304 in HB dataset and 0.486, 0.378 and 0.358 in XZ dataset. Fine tuning enabled the baseline algorithm to perform better on both local datasets, as we observed an increase in the value of DICE and IOU and a decrease in the value of ASD (i.e., 0.593, 0.450 and 0.283 in HB dataset and 0.601, 0.482 and 0.225 in XZ dataset) with corresponding change rate of 33.56%, 29.31% and -6.91% in HB and 23.66%, 27.51% and -37.15% in XZ. Almost all of the change rates are significant except for the -6.91%. Higher values of DICE and IOU, and a lower value of ASD indicate better performance of the segmentation algorithm.

Figure 3 shows examples of segmentation result of the algorithm with and without fine tuning. We observed that the baseline algorithm segmented the lesion region in more details after using the fine tuning method for the HB dataset (i.e., After_HB vs. Undo_HB), which could be reflected by a higher value of ASD that was used to evaluate the algorithms' edge fitting performance. In addition, it is noteworthy that when the baseline algorithm was applied in XZ dataset, a false positive nodule was detected, but after

the algorithm fine tuning the false positive nodule was no longer identified and segmented (i.e., After_XZ vs. Undo_XZ).

3.3 Effect of fine tuning on classification algorithms

As shown in Table 2, the baseline classification algorithm achieved an AUC of 0.881, and the accuracy was 0.846 in the internal testing. When it was applied in two local datasets, the AUC decreased to 0.812 and 0.668, and the accuracy decreased to 0.769 and 0.617 in HB and XZ datasets, respectively. Other metrics of sensitivity, specificity and F1 score also experienced a decreasing tendency in both HB and XZ datasets. However, they exhibited varying degrees of decrease (Figures 4, 5), which is consistent with prior research revealing that the proposed algorithm would display high variability in performance across external datasets (18).

To explore the effect of fine tuning on classification algorithm, the comparison of the validation results between baseline algorithms and fine tuned algorithms, namely MHB and MXZ, was conducted (Table 3). The classification performance of both MHB and MXZ was improved after the fine tuning (Figure 6). Specifically, comparing with the baseline validation results, the MHB had higher AUC (0.851 vs. 0.812), accuracy (0.813 vs. 0.769), sensitivity (0.849 vs. 0.767) and F1 score (0.852 vs. 0.822),

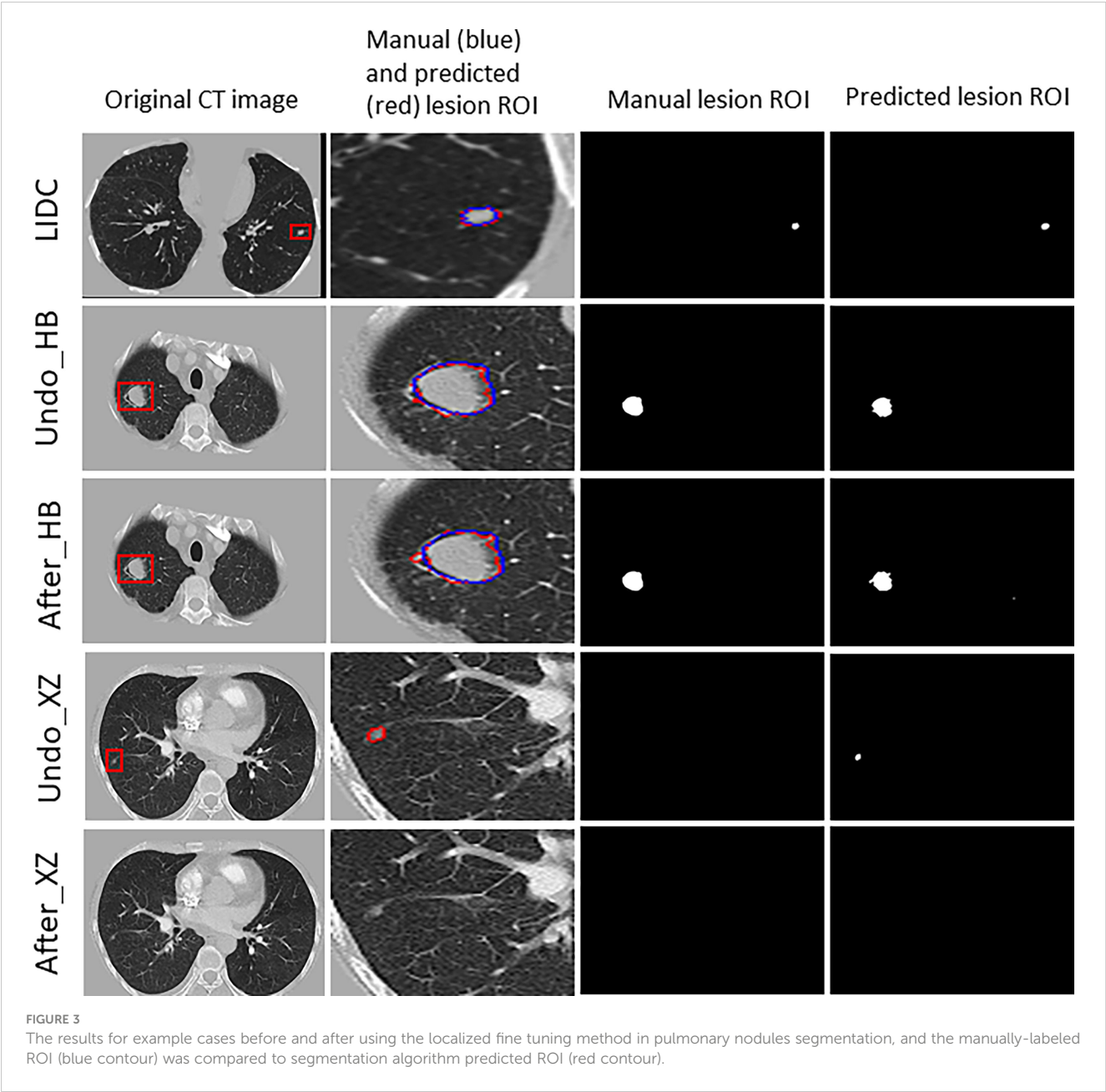
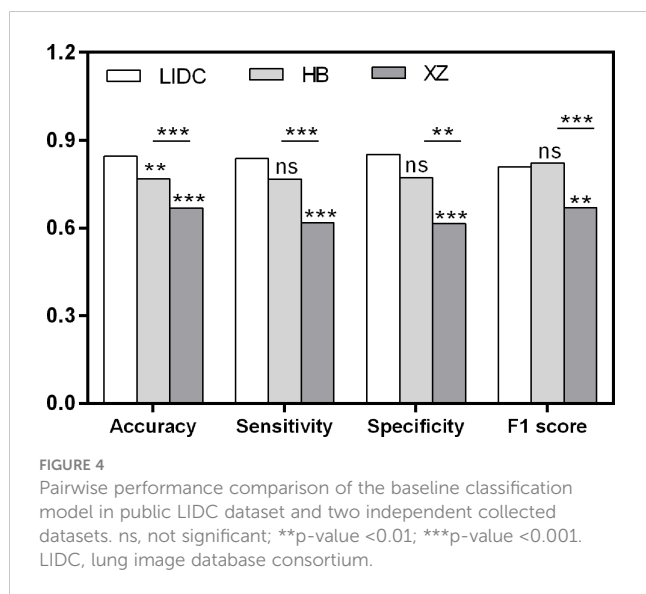


TABLE 2 Performance of baseline classification model in both public dataset and independent collected datasets.

Measure	Performance (95% CI)		
	Datasets		
	LIDC	HB	XZ
AUC	0.881 (0.830-0.920)	0.812 (0.776-0.845)	0.668 (0.615-0.717)
Accuracy	0.846 (0.792-0.888)	0.769 (0.731-0.803)	0.617 (0.565-0.667)
Sensitivity	0.837 (0.744-0.902)	0.767 (0.721-0.808)	0.619 (0.552-0.681)
Specificity	0.852 (0.782-0.903)	0.772 (0.700-0.831)	0.615 (0.530-0.695)
F1 score	0.809 (0.789-0.828)	0.822 (0.803-0.840)	0.668 (0.621-0.713)

AUC, area under the ROC curve.



and their change rate were 4.8%, 5.7%, 10.7% and 3.6%. Though the specificity was slightly decreased by 5.4%, there was no significant difference (0.730 vs. 0.772, $P=0.363$). For MXZ validation results, all the evaluating metrics were increased, including AUC (0.724 vs. 0.668), accuracy (0.696 vs. 0.617), sensitivity (0.684 vs. 0.619), specificity (0.713 vs. 0.615) and F1score (0.737 vs. 0.668), and their change rate were 8.4%, 12.8%, 10.5%, 15.9% and 10.3%.

4 Discussion

In this study, we developed a baseline segmentation algorithm and a baseline classification algorithm with public dataset of LIDC to segment nodules and classify them as being benign or malignant, and then conducted fine tuning for both of them to compare their performance with that of their baseline ones. The results showed that both segmentation and classification process benefit from fine

tuning and end up obtaining higher performance for the site-specific use.

Generally, the development of a computer-aided diagnosis (CAD) scheme consists of the following steps: image preprocessing, ROI segmentation, feature extraction, and finally classification. DL models have been shown to significantly contribute to medical image analysis for the processes of segmentation and classification (21), and many methods have been proposed on optimizing the segmentation and classification algorithm independently (22). Technically, segmentation is used to detect and localize the ROI from the background within the medical image, followed by the segment-based classification task to classify the ROI to a certain class, and the DL model performance may largely rely on the reliable ROI segmentation and good classifier (23). In the current study, we first proposed baseline DL algorithms of segmentation and classification, and compared the performance before and after fine tuning on imaging to explore to what extent the fine tuning can help improve the segmentation and classification process independently.

Algorithms developed on public datasets may not be implied directly on other populations, and rigorous external validation is essential to objectively assess the performance of a detection algorithm (24). In the study, we developed a segmentation and a classification algorithm using public dataset of LIDC, and unlike most of the work with adopted and balanced validation dataset, we applied two external datasets which are inherently different from each other with a significant difference in the distribution of benign and malignant nodules. Thus, the algorithm performance was evaluated in the real-word screening setting, providing objective evidence for the usefulness of the algorithm. It is common to conduct a pilot phase to optimize a triaging threshold of CAD system for external test. However, the threshold choice is balanced between maximal case finding and lower false positive cases without model improvement (25, 26). Therefore, in the study, even with the optimal threshold we observed a decreased performance in the two external tests for both baseline segmentation and classification

TABLE 3 Comparison of the baseline classification model and its fine tuned models.

	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F1 score (95% CI)
Baseline	0.812 (0.776-0.845)	0.769 (0.731-0.803)	0.767 (0.721-0.808)	0.772 (0.700-0.831)	0.822 (0.803-0.840)
MHB	0.851 (0.823-0.875)	0.813 (0.777-0.844)	0.849 (0.809-0.883)	0.730 (0.654-0.791)	0.852 (0.846-0.879)
Rate of change	4.8%	5.7%	10.7%	-5.4%	3.6%
<i>P</i>	0.011	0.080	0.005	0.363	0.030
Baseline	0.668 (0.615-0.717)	0.617 (0.565-0.667)	0.619 (0.552-0.681)	0.615 (0.530-0.695)	0.668 (0.621-0.713)
MXZ	0.724 (0.673-0.770)	0.696 (0.645-0.742)	0.684 (0.619-0.742)	0.713 (0.632-0.786)	0.737 (0.714-0.759)
Rate of change	8.4%	12.8%	10.5%	15.9%	10.3%
<i>P</i>	0.096	0.030	0.157	0.088	0.034

AUC, area under the curve.

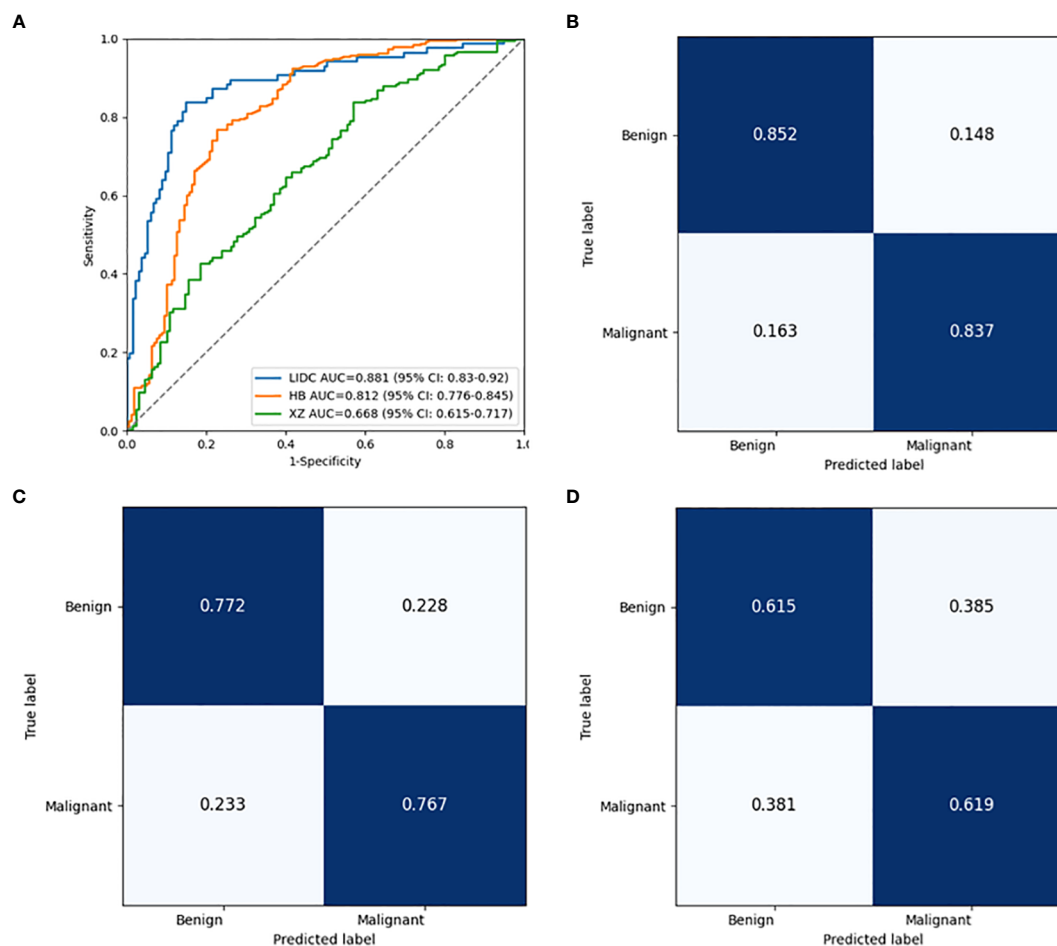


FIGURE 5

Performance comparison of the baseline classification model in public LIDC dataset and two independent collected datasets. (A) The ROC curves in LIDC, HB and XZ datasets. (B) Normalized confusion matrix in LIDC dataset. (C) Normalized confusion matrix in HB dataset. (D) Normalized confusion matrix in XZ dataset. ROC, receiver operating characteristic; LIDC, lung image database consortium.

algorithm (Table 2). The results showed that the algorithm trained by public dataset needs further adjustments for site-specific use, which is consistent with reported research (27, 28).

In previous studies, the deep learning models used for lung nodules segmentation on LIDC dataset obtained the DICE values of over 0.6 (29), and the existed classification algorithm had AUC values of over 0.8 for benign and malignant nodules classification (5, 30, 31), which is similar to our baseline segmentation algorithm and baseline classification algorithm. However, the DICE value decreased when the baseline segmentation algorithm was applied on HB and XZ, and the performance drop could also be detected in the external tests for the baseline classification algorithm. This may result from the significant appearance variances caused by the population and setting differences (32–34). It has been reported that involving multi-center datasets to develop algorithm is effective to keep the algorithm robust to maintain its accuracy across datasets (10, 11). However, it is unclear how many datasets should be exactly included to create a robust detection algorithm to obtain comparable performances of the internal test, especially when those external datasets are significantly different from internal datasets. Furthermore, AlBadawy et al. reported that using multiple institutions for training does not necessarily remove the

dataset shift limitation (32). Model tuning with additional data from specific settings may be an effective way to reduce site-specific biases (11) but few studies revealed its impact on segmentation and classification process alone. In the current study, the baseline models trained by public data set were fine tuned with site-specific images and we observed both segmentation and classification algorithm benefit from the fine tuning, which showed that localized fine tuning would be a potential and well-operated way to develop an automated diagnostic tool to screen lung cancer as both the segmentation process and classification process could get optimized (35). It should be noted that the baseline segmentation algorithm was fine tuned to have as high of a sensitivity as possible for localizing and segmenting the nodules, allowing for false positive reduction, which might be due to that homogeneous features of the local dataset were involved for the learning process.

There are some limitations to this study. First, although both segmentation process and classification process were found improved with the fine tuning, it only focused on lung nodules. For the next step of our study, we aim to expand to other lung abnormality/disease to comprehensively validate the effectiveness of the fine tuning method. Second, the current study was a

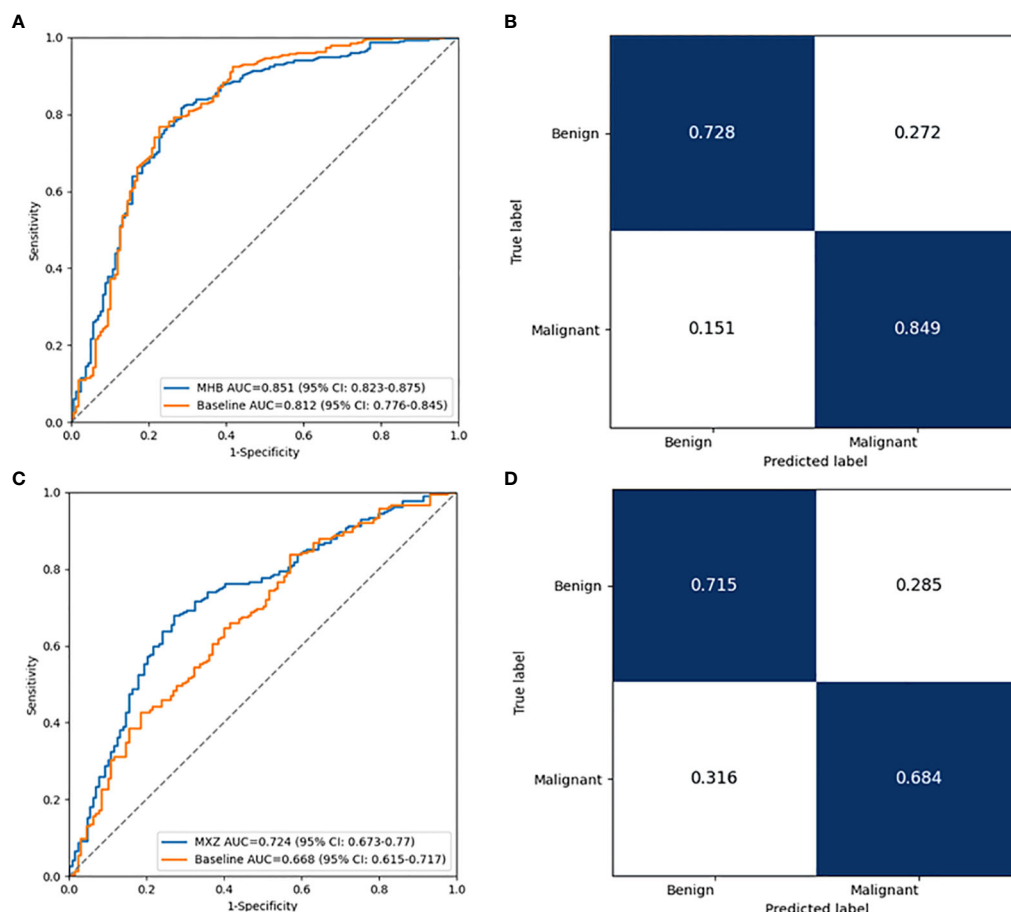


FIGURE 6

Performance comparison between the baseline classification model and fine tuned classification model in two independent collected datasets.

(A) The ROC curves of the two models in HB dataset. (B) Normalized confusion matrix of the fine tuned model in HB dataset. (C) The ROC curves of the two models in XZ dataset. (D) Normalized confusion matrix of the fine tuned model in XZ dataset. ROC, receiver operating characteristic.

retrospective study where both LIDC and two collected datasets were available at the time of study, therefore, a prospective evaluation is needed to further validate the proposed method.

Conclusion

Our work is among the first that conducted the localized fine tuning for DL algorithm on imaging to explore its impact on the segmentation and classification process respectively. Results showed that both segmentation and classification algorithm outperformed their baseline model, which might enable a baseline algorithm be generalized for site-specific use and promote the future in-depth research towards its clinical application.

Data availability statement

The raw data supporting the conclusions of this article will be made available from corresponding author on reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) of the Affiliated Hospital of Hebei University. The informed consent from human participants was waived because this is a retrospective study, and the waiver was indicated in the IRB approval document.

Author contributions

Study design and conception were proposed by JC, LG, LZ and XY. Paper writing was done by JC, LG and LZ. CT scans were collected by JC and LG. AI model training, testing and visualizing were done by LX, LQ and YMFL. All authors interpreted the results and revised the manuscript. All authors read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Program of China (Grant No.: 2019YFE0121400), the Shenzhen Science and Technology Program (Grant No.: KQTD2017033110081833; JSGG20201102162802008; JCYJ20220531093817040), and the Shenzhen Fundamental Research Program (No.: JCYJ20190813153413160), and the Guangzhou Science and Technology Planning Project (No.: 2023A03J0536).

Acknowledgments

We are grateful to the National Cancer Institute and the Foundation for the National Institute of Health, USA on free publicly available LIDC database using in this study.

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
- Toumazis I, Bastani M, Han SS, Plevritis SK. Risk-based lung cancer screening: A systematic review. *Lung Cancer* (2020) 147:154–86. doi: 10.1016/j.lungcan.2020.07.007
- Oudkerk M, Liu S, Heuvelmans MA, Walter JE, Field JK. Lung cancer LDCT screening and mortality reduction — evidence, pitfalls and future perspectives. *Nat Rev Clin Oncol* (2021) 18(3):135–51. doi: 10.1038/s41571-020-00432-6
- Baihua Z, Qi S, Monkam P, Li C, Yang F, Yao Y, et al. Ensemble learners of multiple deep CNNs for pulmonary nodules classification using CT images. *IEEE Access* (2019) 7:110358–71. doi: 10.1109/ACCESS.2019.2933670
- Zhao X, Liu L, Qi S, Teng Y, Li J, Qian W. Agile convolutional neural network for pulmonary nodule classification using CT images. *Int J Comput Assist Radiol Surg* (2018) 13(4):585–95. doi: 10.1007/s11548-017-1696-0
- Gupta A, Saar T, Martens O, Moulec YL. Automatic detection of multisize pulmonary nodules in CT images: Large-scale validation of the false-positive reduction step. *Med Phys* (2018) 45(3):1135–49. doi: 10.1002/mp.12746
- Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology* (2019) 290(1):218–28. doi: 10.1148/radiol.2018180237
- Garau N, Paganelli C, Summers P, Choi W, Alam S, Lu W, et al. External validation of radiomics-based predictive models in low-dose CT screening for early lung cancer diagnosis. *Med Phys* (2020) 47(9):4125–36. doi: 10.1002/mp.14308
- Zhang G, Yang Z, Gong L, Jiang S, Wang L. Classification of benign and malignant lung nodules from CT images based on hybrid features. *Phys Med Biol* (2019) 64(12):125011. doi: 10.1088/1361-6560/ab2544
- Zhou W, Cheng G, Zhang Z, Zhu L, Jaeger S, Lure FYM, et al. Deep learning-based pulmonary tuberculosis automated detection on chest radiography: Large-scale independent testing. *Quant Imag Med Surg* (2022) 12(4):2344–55. doi: 10.21037/qims-21-676
- Kuo P-C, Tsai CC, López DM, Karargyris A, Pollard TJ, Johnson AEW, et al. Recalibration of deep learning models for abnormality detection in smartphone-captured chest radiograph. *NPJ Digit Med* (2021) 4(1):25. doi: 10.1038/s41746-021-00393-9
- Park S, Kim H, Shim E, Hwang B-Y, Kim Y, Lee J-W, et al. Deep learning-based automatic segmentation of mandible and maxilla in multi-center CT images. *Appl Sci* (2022) 12(3):1358. doi: 10.3390/app12031358
- Rundo L, Han C, Nagano Y, Zhang J, Hataya R, Militello C, et al. USE-net: Incorporating squeeze-and-excitation blocks into U-net for prostate zonal segmentation of multi-institutional MRI datasets. *Neurocomputing* (2019) 365:31–43. doi: 10.1016/j.neucom.2019.07.006
- Singh H, Mhasawade V and CR. Generalizability challenges of mortality risk prediction models: A retrospective analysis on a multi-center database. *PLoS Digit Health* (2022) 1(4):e0000023. doi: 10.1371/journal.pdig.0000023
- Winter A, Aberle DR, Hsu W. External validation and recalibration of the Brock model to predict probability of cancer in pulmonary nodules using NLST data. *Thorax* (2019) 74(6):551. doi: 10.1136/thoraxjnl-2018-212413

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors LG, LX, and LQ are employed by the company Shenzhen Zhiying Medical Imaging. Author YMFL is a stockholder of the company Shenzhen Zhiying Medical Imaging.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ensor J, Snell KIE, Debray TPA, Lambert PC, Look MP, Mamas MA, et al. Individual participant data meta-analysis for external validation, recalibration, and updating of a flexible parametric prognostic model. *Stat Med* (2021) 40(13):3066–84. doi: 10.1002/sim.8959
- Steyerberg EW, Borsboom GJJM, van Houwelingen HC, Eijkemans MJC, Habbema JDF. Validation and updating of predictive logistic regression models: A study on sample size and shrinkage. *Stat Med* (2004) 23(16):2567–86. doi: 10.1002/sim.1844
- Janssen KJM, Moons KGM, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* (2008) 61(1):76–86. doi: 10.1016/j.jclinepi.2007.04.018
- Armato III SG, McLennan G, McNitt-Gray MF, Meyer CR, Yankelevitz D, Aberle DR, et al. Lung image database consortium: Developing a resource for the medical imaging research community. *Radiology* (2004) 232(3):739–48. doi: 10.1148/radiol.2323032035
- Zhang B, Qi S, Wu Y, Pan X, Yao Y, Qian W, et al. Multi-scale segmentation squeeze-and-excitation UNet with conditional random field for segmenting lung tumor from CT images. *Comput Methods Programs Biomed* (2022) 222:106946. doi: 10.1016/j.cmpb.2022.106946
- Bibi A, Khan M, Javed M, Tariq U, Kang B-G, Nam Y, et al. Skin lesion segmentation and classification using conventional and deep learning based framework. *CMC Comput Mater Con* (2022) 71(2):2477–95. doi: 10.32604/cmc.2022.018917
- Wang X, Jiang L, Li L, Xu M, Deng X, Dai L, et al. Joint learning of 3D lesion segmentation and classification for explainable COVID-19 diagnosis. *IEEE Trans Med Imaging* (2021) 40(9):2463–76. doi: 10.1109/tmi.2021.3079709
- Dalila F, Zohra A, Reda K, Hocine C. Segmentation and classification of melanoma and benign skin lesions. *Optik* (2017) 140:749–61. doi: 10.1016/j.ijleo.2017.04.084
- Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med* (2015) 162(1):W1–W73. doi: 10.7326/M14-0698
- Fehr J, Konigorski S, Olivier S, Gunda R, Surujdeen A, Gareta D, et al. Computer-aided interpretation of chest radiography reveals the spectrum of tuberculosis in rural south Africa. *NPJ Digit Med* (2021) 4(1):106. doi: 10.1038/s41746-021-00471-y
- Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and validation of a deep learning-based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs. *Clin Infect Dis* (2019) 69(5):739–47. doi: 10.1093/cid/ciy967
- Nam JG, Park S, Hwang EJ, Lee JH, Jin K-N, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology* (2018) 290(1):218–28. doi: 10.1148/radiol.2018180237
- Wang X, Liang G, Zhang Y, Blanton H, Bessinger Z, Jacobs N. Inconsistent performance of deep learning models on mammogram classification. *J Am Coll Radiol* (2020) 17(6):796–803. doi: 10.1016/j.jacr.2020.01.006
- Liu H, Cao H, Song E, Ma G, Xu X, Jin R, et al. A cascaded dual-pathway residual network for lung nodule segmentation in CT images. *Phys Med* (2019) 63:112–21. doi: 10.1016/j.ejmp.2019.06.003

30. Xie Y, Zhang J, Xia Y, Fulham M, Zhang Y. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. *Inform Fusion* (2018) 42:102–10. doi: 10.1016/j.inffus.2017.10.005
31. Sun W, Zheng B, Qian W. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. *Comput Biol Med* (2017) 89:530–9. doi: 10.1016/j.compbiomed.2017.04.006
32. AlBadawy EA, Saha A, Mazurowski MA. Deep learning for segmentation of brain tumors: Impact of cross-institutional training and testing. *Med Phys* (2018) 45 (3):1150–8. doi: 10.1002/mp.12752
33. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. *N Engl J Med* (2021) 385(3):283–6. doi: 10.1056/NEJMc2104626
34. Torralba A, Efros AA. Unbiased look at dataset bias. : *CVPR 2011*. (2011). 1521–8 doi: 10.1109/CVPR.2011.5995347
35. Ozdemir O, Russell RL, Berlin AA. A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. *IEEE Trans Med Imaging* (2020) 39(5):1419–29. doi: 10.1109/tmi.2019.2947595



OPEN ACCESS

EDITED BY

Laura Curiel,
University of Calgary, Canada

REVIEWED BY

Justyna Chatubiriska-Fendler,
Military Institute of Medicine, Poland
Jung Hun Oh,
Memorial Sloan Kettering Cancer Center,
United States

*CORRESPONDENCE

Wenbo Meng
✉ mengwb@lzu.edu.cn

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 27 December 2022

ACCEPTED 10 March 2023

PUBLISHED 29 March 2023

CITATION

Hu W, Zhang X, Saber A, Cai Q, Wei M,
Wang M, Da Z, Han B, Meng W and Li X
(2023) Development and validation of a
nomogram model for lung cancer based
on radiomics artificial intelligence score
and clinical blood test data.
Front. Oncol. 13:1132514.
doi: 10.3389/fonc.2023.1132514

COPYRIGHT

© 2023 Hu, Zhang, Saber, Cai, Wei, Wang,
Da, Han, Meng and Li. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Development and validation of a nomogram model for lung cancer based on radiomics artificial intelligence score and clinical blood test data

Wenteng Hu^{1,2}, Xu Zhang¹, Ali Saber³, Qianqian Cai¹, Min Wei^{1,4},
Mingyuan Wang^{1,5}, Zijian Da¹, Biao Han^{1,2}, Wenbo Meng^{1,6*}
and Xun Li^{1,6}

¹The First Clinical Medical School of Lanzhou University, Lanzhou, Gansu, China, ²Department of Thoracic Surgery, The First Hospital of Lanzhou University, Lanzhou, Gansu, China, ³Saber Medical Genetics Laboratory, Almas Medical Complex, Rasht, Iran, ⁴Department of Emergency, The First Hospital of Lanzhou University, Lanzhou, Gansu, China, ⁵Department of Ultrasonography, The First Hospital of Lanzhou University, Lanzhou, Gansu, China, ⁶Department of General Surgery, The First Hospital of Lanzhou University, Lanzhou, Gansu, China

Background: Artificial intelligence (AI) discrimination models using single radioactive variables in recognition algorithms of lung nodules cannot predict lung cancer accurately. Hence, we developed a clinical model that combines AI with blood test variables to predict lung cancer.

Methods: Between 2018 and 2021, 584 individuals (358 patients with lung cancer and 226 individuals with lung nodules other than cancer as control) were enrolled prospectively. Machine learning algorithms including lasso regression and random forest (RF) were used to select variables from blood test data, Logistic regression analysis was used to reconfirm the features to build the nomogram model. The predictive performance was assessed by performing the receiver operating characteristic (ROC) curve analysis as well as calibration, clinical decision and impact curves. A cohort of 48 patients was used to independently validate the model. The subgroup application was analyzed by pathological diagnosis.

Findings: A total of 584 patients were enrolled (358 lung cancers, 61.30%, 226 patients for the control group) to establish the model. The integrated model identified eight potential factors including carcinoembryonic antigen (CEA), AI score, Pro-Gastrin Releasing Peptide (ProGRP), cytokeratin 19 fragment antigen 21-1 (CYFRA211), squamous cell carcinoma antigen (SCC), indirect bilirubin (IBIL), activated partial thromboplastin time (APTT) and age. The area under the curve (AUC) of the nomogram was 0.907 (95% CI, 0.881–0.929). The decision and clinical impact curves showed good predictive accuracy of the model. An AUC of 0.844 (95% CI, 0.710–0.932) was obtained for the external validation group.

Conclusion: The nomogram model integrating AI and clinical data can accurately predict lung cancer, especially for the squamous cell carcinoma subtype.

KEYWORDS

lung cancer, artificial intelligence, prediction model, pulmonary nodule, machine learning (ML)

1 Introduction

Lung cancer is the leading cause of cancer mortality worldwide now. But patients often have a long course without atypical symptoms and signs (1). Therefore, early diagnosis is not possible in most cases, and 5-year survival rate is only 16.1% (2). Low-dose computed tomography (LDCT) is the main method for public physical screening. The tumor markers assessment in hospital including carcinoembryonic antigen (CEA) and cytokeratin 19 fragment antigen21-1(CYFRA21-1) can improve the diagnosis rate (3). Artificial intelligence (AI) models are a step forward from automated nodule diagnosis, as they typically do not require nodule measurement or data entry.

Available radioactive prediction models include the Mayo model, Veterans Administration (VA) model, Brock University model, and Peking University People's Hospital model (PKUPH). However, these models mainly focus on the CT performance of pulmonary nodules and currently, but there is not any model integrating routine blood test data, pathological data and the radioactive models combined with pathological data for accurate prediction of lung cancer (4). In this study, we aimed to build an integrated prediction model for pulmonary nodule diagnosis based on clinical laboratory data and the VA model (5). Thus, we developed a nomogram model incorporating pathological-based subgroup analysis as a timely and efficient tool for clinical application (6).

2 Methods

2.1 Training population and study design

This retrospective study was conducted in the thoracic surgery department of the First Hospital of Lanzhou University in China, following the Declaration of Helsinki, and was approved by the Ethical Committee of the First Hospital of Lanzhou University (reference number: LDYYLL2021-257). Written informed consent was obtained from all patients. The principles of this study are followed with TRIPOD (The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis).

2.2 Patients enrollment

Consecutive pulmonary nodule patients who got AI tool assessment before surgery Between January 2018 to December 2021 were included in this study. Tumor pathological subtypes were assessed by an experienced pathologist. Exclusion criteria were: having accurate pathological data, multiple metastatic tumors, cases with missing data, lung transplant or previous history of lung surgery, and having radiotherapy and chemotherapy.

A total of 861 eligible patients were screened initially. Among these, 142 patients with current clear pathological diagnosis before the operation, 58 multiple metastatic tumors, 27 cases with missing data, and 50 patients with a history of lung surgery, radiotherapy, or chemotherapy were excluded from further analysis. Finally, 584 eligible patients were included in the study to train the model. [Figure 1](#) shows the flowchart for patient recruitment in this study. In addition, a total of 48 eligible patients from January 2022 to May 2022 were recruited to validate the predictive value of the model.

2.3 Data collection

Data was collected independently by two reviewers (WT. H. and X. Z.). More than 76 pre-surgical parameters were evaluated in the developing model: 1. Patients' basic line characters include age, gender, history of hypertension, diabetes, history of smoking, drinks, and family history. 2. Blood laboratory tests including all parameters of coagulation and blood routine examination data. 3. All routine blood biochemical function test parameters. 4. The lung cancer-associated serum tumor markers include CEA, CYFRA 21-1, squamous cell carcinoma antigen(SCC), neuron-specific enolase (NSE), and Ferritin(FER). 5. The AI assessment score from σ -Discover/Lung Nodule intelligent diagnosis system (the system has got permission from the Chinese Medical Association) with a sensitivity of 80.17% and a specificity of 70.35%. 6. The pathological data include lung cancer subtype, the degree of tumor differentiation, tumor infiltration, the tumor node metastasis (TNM) stage (7).

3 Statistical analysis and development of a nomogram

3.1 Prediction model development

First, potential risk factors were identified through machine learning methods from the routine blood test data. The selected variables, AI score and patients' basic line character were used as candidate parameters for model development. Then, the risk variables selection was calculated by stepwise multivariate logistic regression (backward, $p < 0.05$) (8). In addition, different histopathological subtype were analyzed to verify the accuracy of the model for the recognition of different lung cancer subtypes. For clinical application, a nomogram figure was established as an integrated clinical prediction tool.

3.2 Machine learning model for variable selection

The machine learning methods were implemented through R (version 4.1.1). The machine learning methods including Least absolute shrinkage and selection operator (LASSO) regression and random forest (RF) were used to identify important features. Lasso regression can handle the multicollinearity problem of the available features and RF enables the implementation of variable selection procedures based on their impact on outcome prediction. RF parameters were optimized in logarithmic steps around their default values (using 500 trees, and a random subspace with dimensionality equal to the rounded value of the square of the number of features). Ten-fold cross-validation and external test set validation were both employed to validate the reliability of the model (9).

3.3 Logistic regression model method

Data were analyzed using SPSS v.22.0 (IBM, Armonk, New York, USA). Patients were grouped by postoperative pathological diagnosis. Univariate and multivariate logistic regressions classified the risk factors for lung cancer. The regression models either used chi-square test or student's t-test for patients' basic features (age, sex, etc.) analysis. P-values below 0.05 were considered statistically significant. Adjusted odds ratios (ORs) and corresponding 95% confidence intervals (95% CIs) were calculated. The Hosmer-Lemeshow test was used to assess the fitness of the model. The accuracy of the model was checked by plotting DCA (decision curve analysis) curve and CIC (clinical impact curve) curves using predicted probabilities against the actual probabilities. The receiver operating characteristic (ROC) and the area under the curve (AUC) were estimated for discrimination (7).

3.4 Subgroup analysis

After model construction, further subgroups analysis according to the postoperative pathological diagnosis. For patients'

pathological diagnosis types including squamous cell carcinoma (SQCC), adenocarcinoma (AD), other tumors such as non-small cell lung cancer (NSCLC), and neuroendocrine tumors patients, The predictive ability of different pathological types of cancer was tested using the integrated model, and the difference between the simple AI predictor and the nomogram was also compared. The prediction performance was estimated by De long test for the AUC, decision curve, and clinical impact curve (10).

3.5 External validation

An independent external validation from January 2022 to May 2022 in the First Hospital of Lanzhou University was performed by using the nomogram according to the cut-off parameter. The ROC curve, DCA curve, and CIC curve analysis were performed to validate the accuracy of the model by estimating the difference between the integrated nomogram from the modeling cohort and validation set.

36 Role of funding source

The funders had no involvement in study design, data collection, data analysis, interpretation of findings, the writing of this paper, or the decision to submit the paper for publication. There was no commercial support. The corresponding author (WBM) had full access to all the data in the study and had final responsibility for the decision to submit it for publication.

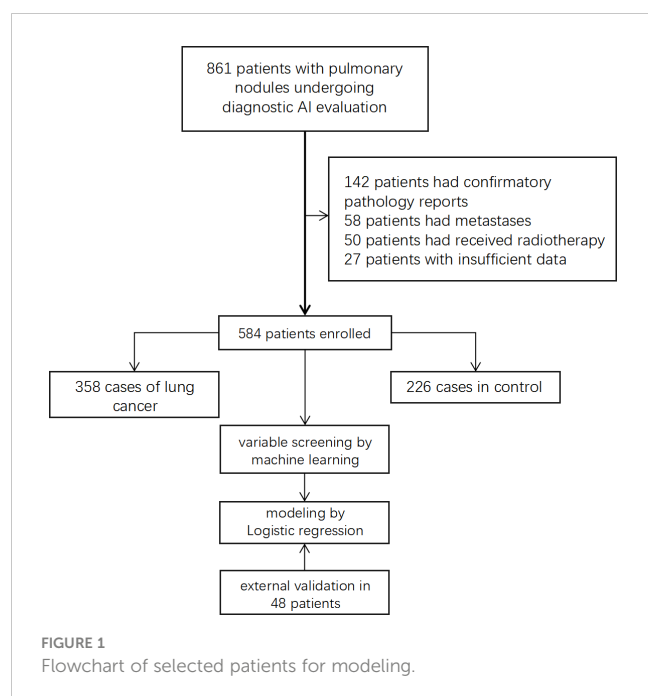
4 Results

4.1 Patient characteristic

A total of 584 eligible patients were enrolled (Figure 1). All patients' basic line characteristics were analyzed before modeling. There was no difference between lung cancer patients and control group in sex, hypertension, diabetes, family history, drinking, and history of chronic pulmonary diseases before surgery. However, patients with age and smoking history before surgery had a higher rate of carcinoma (Table 1).

4.2 Variables selection

We used the LASSO algorithm to select feature variables from laboratory test data. Except for uric acid (UA), all 76 variables excluded collinearity and could be included in the variable selection using the RF method (Figure 2). To obtain the best set of features, the importance of each variable was calculated; 30 features were identified by the RF method finally. These steps were performed by the "RandomForest" package that has been illustrated in Figure 3.



4.3 Prediction model by logistic analysis

Factors found by the RF algorithms, AI score and baseline data were calculated in the univariate and multivariate logistic regression analysis in training group.

For lung cancer patients, radiological AI score higher than 77 (OR=1.098; 95% CI, 1.074-1.123), serum levels of CEA higher than 2.3 µg/L (OR=1.193; 95% CI, 1.019-1.396), serum levels of ProGRP higher than 40.2 µg/L (OR=1.014; 95% CI, 1.001-1.028), serum levels of CYFRA211 higher than 2.5 µg/L (OR=1.714; 95% CI, 1.356-2.167), serum levels of SCC higher than 0.8 U/L

(OR=2.336; 95% CI, 1.240-4.402), serum levels of IBIL higher than 16.8 U/L (OR=1.057; 95% CI, 1.009-1.107), the APTT shorter than 34 s (OR=0.916; 95% CI, 0.862-0.974) and age>52 (OR=1.045; 95% CI, 1.018-1.072) were high risk predictors for developing integrated model in lung nodules patients (Tables 2, 3). The AUC of the nomogram for the prediction of lung cancer was 0.907 (95% CI, 0.881-0.929). The De Long test for comparing the performance of the integrated model and the AI model evaluation was 0.001 and accuracy was examined by DCA and CIC curve analysis (Figure 4).

4.4 External validation of the model

A cohort of 48 patients (32 lung cancer, 66.7%, 16 patients for the control group) was included to validate the nomogram by the cut-off value from the training set, followed by the ROC, DCA, and CIC curve analyses. The prediction ROC curve with an AUC of 0.844 (95% CI, 0.710-0.932) with a sensitivity of 81.20% and a specificity of 87.50%, calibration, DCA and CIC curves showed that the accuracy is in the fitting range. The external cohort showed that our integrated model is in line with the clinical setting (Figure 5).

4.5 Nomogram of the model

To facilitate the application of our model, we established an open access nomogram prediction tool. Users could predict pulmonary nodules by 6 features combined AI scores in the figure (Figure 6). Each factor has a prediction reference value based on OR which shows the weight of each parameter and a total score will distinguish between healthy individuals and patients with lung cancer.

TABLE 1 All patients' baseline clinical features analysis.

	Lung cancer (N = 358)	Non-cancer (N = 226)	χ^2/Z	P
Age (year)	57.65 ± 9.01	48.82 ± 12.73	24.523	0.000
Sex (male), n (%)	205(57.26%)	121(53.54%)	0.779	0.378
Hypertension	57(15.92%)	23(10.18%)	3.260	0.071
Diabetes	38(10.61%)	15(6.64%)	2.656	0.103
Smoking history	114(31.84%)	37(16.37%)	20.8	0.001
Drink history	84(23.46%)	37(16.37%)	1.950	0.153
Family history*	17(4.75%)	4(1.77%)	3.540	0.060
Chronic diseases	108(30.17%)	53(23.45%)	3.013	0.077
Nodule status				
Multiple	126(35.20%)	72(31.86%)	0.688	0.407
Burr around nodule	148(41.34%)	76(33.63%)	3.458	0.062
Nodular calcification	82(22.91%)	34(15.04%)	5.606	0.108
Pleural traction	42(11.73%)	16(7.08%)	3.352	0.067
Solid nodule	125(34.91%)	62(27.43%)	3.824	0.051

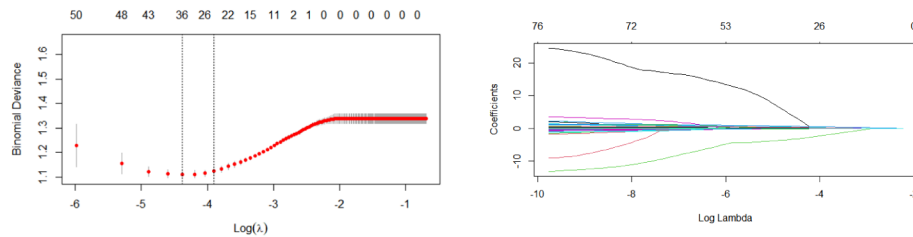


FIGURE 2

Lasso regression for variable selection.

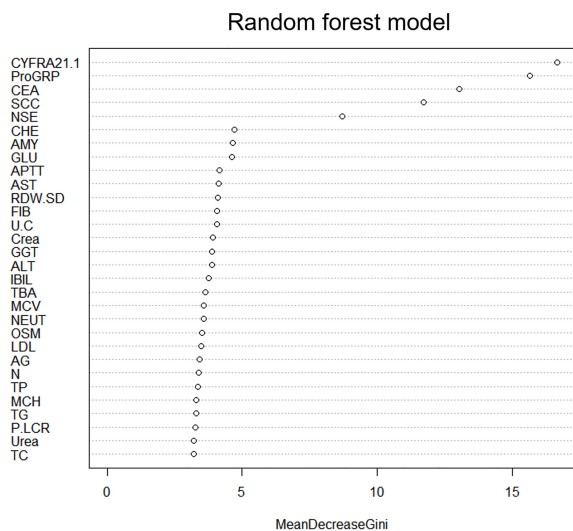


FIGURE 3

Performance of top-ranking variables selected by RF.

4.6 Subgroup analysis

The nomogram was analyzed in different histological subtypes of lung cancer. For patients with SQCC, the integrated nomogram model showed a better predictive performance with an AUC of 0.827 (95% CI 0.794–0.857) as compared to the AI model achieved an AUC of 0.668 (95% CI 0.628–0.707). The p-value of the De Long test was 0.001 and accuracy was examined by DCA and CIC curve analysis, which showed the integrated model has a more accurate prediction ability. For patients with AD, the integrated nomogram model showed slightly better predictive performance with an AUC of 0.799 (95% CI 0.764–0.831) as compared to the AI model with an AUC of 0.735 (95% CI 0.697–0.770). The p-value of the De Long test is 0.001 and accuracy was examined by DCA and CIC curve analysis, which showed the integrated model has a more accurate prediction accuracy. For patients with other types of lung tumors, our integrated nomogram model showed no difference in predictive performance with an AUC of 0.728 (95% CI 0.690–0.764) in comparison with the AI model with an AUC of 0.553 (95% CI 0.491–0.574). The p-value of the De Long test was 0.052, the accuracy test in the DCA and CIC curve also proved there is no

significant predictive differentiation for the integrated model (Appendix Figures 1–3).

5 Discussion

Lung cancer is the main cause of cancer-related deaths worldwide. Early diagnosis can facilitate intraoperative planning procedures (11). Several risk factors such as age, gender, imaging signs of nodules, and tumor markers are related to the malignancy of pulmonary nodules (12). With the development of artificial intelligence technology, the machine learning models provided a better alternative for creating applicable predictive clinical diagnosis tools. In this study, we developed and validated a diagnostic nomogram model to improve the diagnostic accuracy of lung cancer based on AI tools and clinical data (3, 10, 13).

The integrated model can strongly discriminate between lung malignancies and other pulmonary nodules. The model has the AUC of 0.907, sensitivity of 88.2%, and specificity of 85.3%. In addition, the p-value of the Hosmer Lemeshow test was 0.919, respectively, and the p-value of the De long test with AI was 0.001. The parameters and DCA, CIC, and calibration curve analyses revealed that our integrated model has an excellent predictive accuracy as compared to the AI model only.

The subgroup analysis for different histopathology subtypes demonstrated that for the SQCC, and AD, the integrated nomogram has a more accurate predictive performance advantage compared to the AI. External validation also proved that the integrated model has a better predictive value. We established a convenient and accurate prediction nomogram tool that could be utilized in the clinical setting.

AI systems based on radiomics features, calculated based on the LDCT images, are widely used for the screening and diagnosis of lung cancer. Current studies support the use of AI prediction models as an effective approach for early diagnosis of lung cancer (14). In our study, when the AI assessment score is higher than 77 by the AI system, the risk of lung cancer will increase and the OR has been applied in the nomogram. AI system scoring is the baseline step in this integrated model (15).

Serum tumor markers in serum have great diagnostic value for preoperative diagnosis. CEA, ProGRP, CYFRA211, and SCC can be used in detecting lung cancer; hence, it is necessary to combine serum tumor markers to improve the diagnostic accuracy (6). The

TABLE 2 Univariate and multivariate logistic regressions of risk factors for lung cancer.

n/N	Univariate analysis		Multivariate analysis	
	OR (95% CI)	p value	OR (95% CI)	p value
Age (year)	1.080(1.061–1.100)	0.000	1.045(1.018–1.072)	0.001
Smoking history	2.387(1.573–3.620)	0.001		
Hypertension	2.238(1.317–4.115)	0.004		
AI levels	1.099(1.079–1.0119)	0.000	1.098(1.074–1.123)	0.000
CEA	1.516(1.331–1.727)	0.000	1.193(1.019–1.396)	0.028
Cyfra211	1.906(1.607–2.259)	0.001	1.714(1.356–2.167)	0.000
ProGRP	1.033(1.021–1.044)	0.003	1.014(1.001–1.028)	0.038
SCC	4.904(2.882–8.342)	0.000	2.336(1.240–4.402)	0.009
NSE	1.072(1.035–1.110)	0.000		
CHE	0.973(0.883–1.073)	0.588		
AMY	1.008(1.001–1.015)	0.017		
GLU	1.074(0.957–1.204)	0.224		
APTT	0.946(0.909–0.984)	0.006	0.916(0.862–0.974)	0.005
AST	0.978(0.946–0.992)	0.002		
RDWSD	1.045(1.003–1.089)	0.037		
FIB	1.389(1.125–1.715)	0.002		
U/C	0.993(0.986–1.001)	0.081		
Crea	1.026(1.012–1.041)	0.010		
GGT	0.997(0.993–1.001)	0.181		
ALT	0.987(0.978–0.995)	0.003		
IBIL	1.040(1.010–1.071)	0.010	1.057(1.009–1.107)	0.019
TBA	1.001(0.968–1.035)	0.956		
MCV	1.045(1.013–1.078)	0.005		
OSM	1.059(1.021–1.098)	0.002		
LDL	1.384(1.101–1.740)	0.005		
AG	1.038(0.979–1.101)	0.211		
TP	1.000(0.991–1.008)	0.918		
MCH	1.109(1.026–1.199)	0.009		
TG	0.946(0.851–1.051)	0.299		
P/LCR	1.002(0.987–1.016)	0.839		
Urea	1.050(0.935–1.179)	0.041		
TC	1.219(1.031–1.440)	0.020		
N%	1.015(0.997–1.033)	0.103		
NEUT	1.057(0.944–1.184)	0.335		

TABLE 3 The predicted value parameters of risk factors.

Factors	Cut-off	sensitivity	specificity	AUC
Age (year)	52	72.91%	58.41%	0.710
AI levels	77	80.17%	70.35%	0.819
CEA	2.3 µg/L	60.6%	72.1%	0.689
Cyfra211	2.5 µg/L	56.98%	73.34%	0.715
ProGRP	40.2 µg/L	66.76%	64.60%	0.691
SCC	0.8 U/L	55.03%	74.78%	0.703
APTT	34 s	69.0%	42.10%	0.557
IBIL	16.8 U/L	27.6%	94.25%	0.551

integrated model took the application of serum tumor markers more convenient for patients as they are substantially less surgical and cost-effective than other methods. This is consistent with the clinical practice of serum tumor markers angiogenesis and neovascularization in malignancy cancer. Furthermore, the serum tumor markers hold a large proportion in our nomogram. It also suggests that we need to pay more attention to serum tumor markers as prediction parameters for lung cancer (16).

Age is one of the common risk factors for tumor course. In this study, age was positively associated with lung cancer, the optimal cutoff value was 52 years old. Thus, physical examination and screening are necessary for the prevention of lung cancer in the elderly.

Previous studies have shown that indirect bilirubin (IBIL) levels had an influence on survival times in 1,617 patients with resectable lung cancer (17). The optimal cutoff value for serum IBIL was 2.5 µg/L with a sensitivity of 27.6% and specificity of 94.25% suggesting that IBLB inhibits the mTOR pathway by altering the activity of the AMPK pathway leading to lung cancer metastasis. Moreover, APTT is one of the routine indexes of hemostatic examination for surgical patients, the optimal cutoff value for APTT time was 34 S with a sensitivity of 60.6% and a specificity of 72.12% (18). Tumor cell products presented on their surface or substances secreted in the microenvironment may either directly trigger clotting system activation or indirectly trigger it by stimulating extravascular host

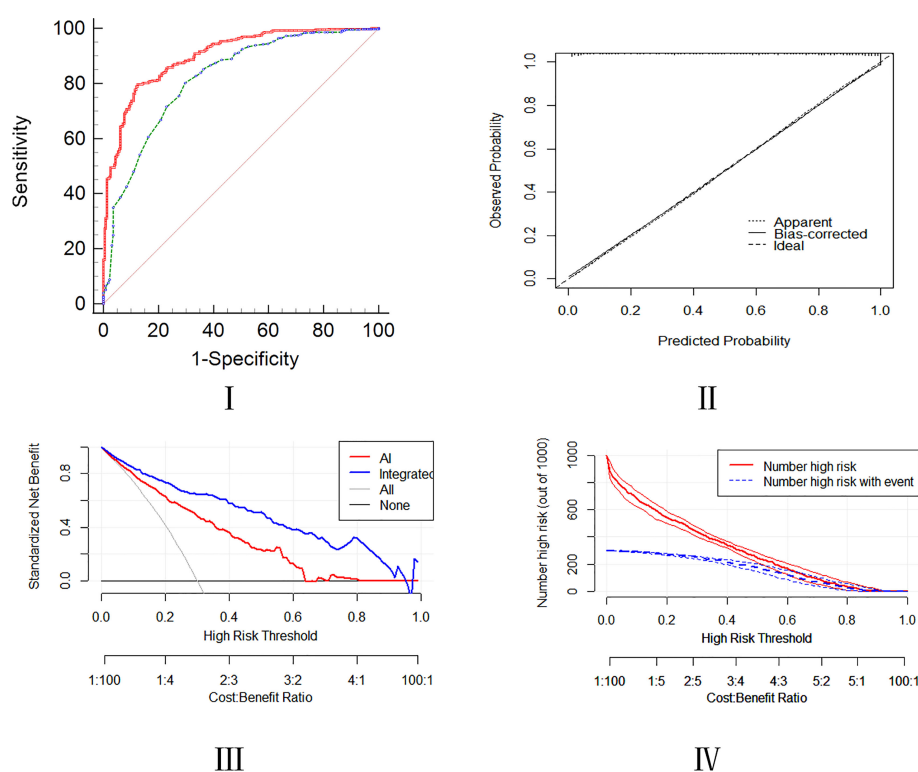


FIGURE 4 Assessment of predictive ability of the integrated model using the ROC, calibration, DCA, and CIC curves.

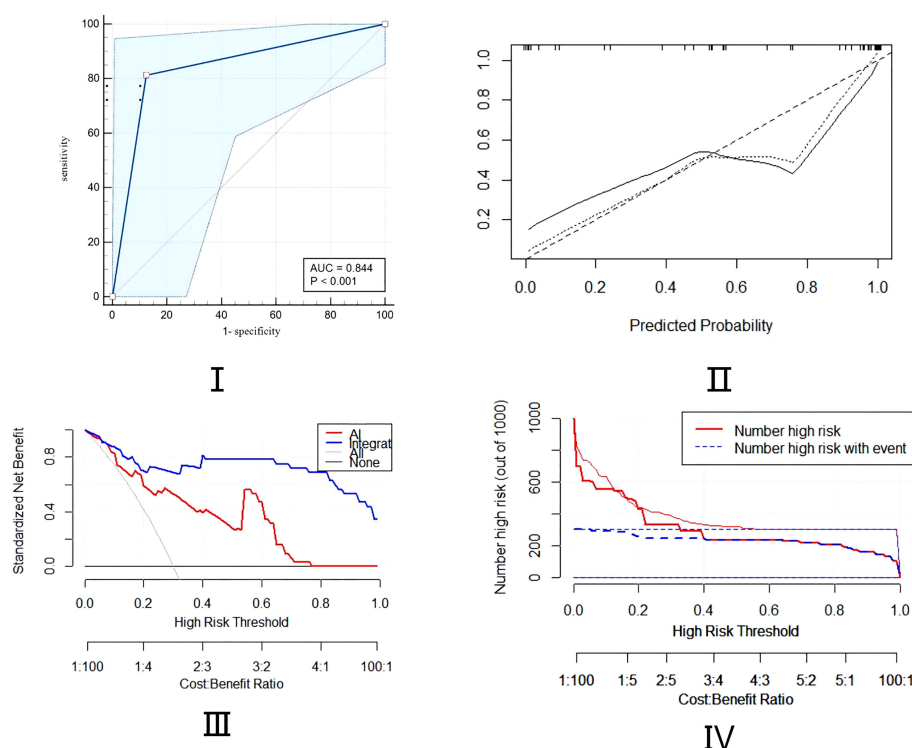


FIGURE 5

Assessment of predictive ability of the integrated model in external validation cohort using ROC, calibration, DCA, and CIC curves.

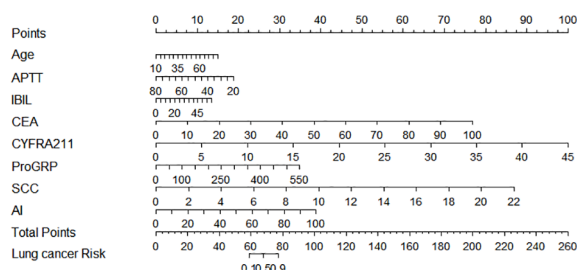


FIGURE 6

The application of the integrated nomogram model.

cells to release procoagulant products (19). This can be also one of the potential targets for cancer detection in the future.

To our knowledge, this is the first diagnostic integrated nomogram model combined with AI tool and clinical blood test data for lung cancer. The validated nomogram showed a high predictive value through the calibration and accuracy test. By the nomogram, the AUC for 8 variables for lung cancer prediction was 0.907 (95% CI, 0.881-0.929), and the p-value of the De Long test is 0.001, which is superior to any single radionics prediction model.

There are still some limitations in our study. First, the nomogram only suits those lung nodule patients instead of routine physical examination for the general population. Secondly, our findings were based on a single-center retrospective study of the eastern Asian population, with an inherent bias with missing data. For future model validation and correction,

prospective global multicenter validation and large-scale studies are needed (20).

In conclusion, CEA, AI score, serum ProGRP, CYFRA211, SCC, IBIL, APTT, and age are potential independent factors that can be used for diagnosis of lung cancer. The presented nomogram, as a less invasive and convenient approach, can accurately predict lung cancer in patients with lung nodules, especially for the SQCC subtype to avoid unnecessary surgical resection.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#). Further inquiries can be directed to the corresponding author.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

The list of full authors in this study is as follows: WH, XZ, AS, QC, MW, MYW, ZD, BH, WM, and XL. Correspondence author:

WM, WH, XZ, AS, and WM: protocol development, drafting of this manuscript, critical revision of the manuscript for significant intellectual content. WH, XZ, QC, MW, MYW, ZD, and BH: conducted clinical trials, patient enrollment, and acquired data. WH, BH, XL, and WM: providing personnel, environmental support, and tools and instruments that are vital for the project. WH, XZ, and WM: taking responsibility for statistical analysis, logical interpretation, and presentation of the results. WH, XZ, and WM: taking responsible for pathology and figures. WH, XZ, AS, QC, MW, BH, and WM: Review the article before submission not only for spelling and grammar but also for its intellectual content. WH, XZ, WM, and XL: constructing an idea or hypothesis for the manuscript, providing critical revision. WH, XZ, and WM accessed and were responsible for the raw data and the models in the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Youth Science and Technology Fund of Gansu Province (18JR3RA305).

References

- Mao Y, Yang D, He J, Krasna MJ. Epidemiology of lung cancer. *Surg Oncol Clinics North America* (2016) 25(3):439–45. doi: 10.1016/j.soc.2016.02.001
- Nasim F, Sabath BF, Eapen GA. Lung cancer. *Med Clinics North America* (2019) 103(3):463–73. doi: 10.1016/j.mcna.2018.12.006
- Chen K, Nie Y, Park S, Zhang K, Zhang Y, Liu Y, et al. Development and validation of machine learning-based model for the prediction of malignancy in multiple pulmonary nodules: Analysis from multicentric cohorts. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2021) 27(8):2255–65. doi: 10.1158/1078-0432.CCR-20-4007
- Raghu VK, Zhao W, Pu J, Leader JK, Wang R, Herman J, et al. Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. *Thorax* (2019) 74(7):643–9. doi: 10.1136/thoraxjnl-2018-212638
- Toumazis I, Bastani M, Han SS, Plevritis SK. Risk-based lung cancer screening: A systematic review. *Lung Cancer* (2020) 147:154–86. doi: 10.1016/j.lungcan.2020.07.007
- Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, et al. Biomarkers in lung cancer screening: Achievements, promises, and challenges. *J Thorac Oncol* (2019) 14(3):343–57. doi: 10.1016/j.jtho.2018.11.023
- Zhou T, Zhu C, Shi F. Application of radiomics in classification and prediction of benign and malignant lung tumors. *Chin J Med instrumentation* (2020) 44(2):113–7.
- Zhang X, Yue P, Zhang J, Yang M, Chen J, Zhang B, et al. A novel machine learning model and a public online prediction platform for prediction of post-ERCP-cholecystitis (PEC). *eClinicalMedicine* (2022) 48:101431. doi: 10.1016/j.eclim.2022.101431
- Wu J, Bai J, Wang W, Xi L, Zhang P, Lan J, et al. ATBdiscrimination: An in silico tool for identification of active tuberculosis disease based on routine blood test and T-SPOT.TB detection results. *J Chem Inf Model* (2019) 59(11):4561–8. doi: 10.1021/acs.jcim.9b00678
- She Y, Jin Z, Wu J, Deng J, Zhang L, Su H, et al. Development and validation of a deep learning model for non-small cell lung cancer survival. *JAMA Netw Open* (2020) 3(6):e205842. doi: 10.1001/jamanetworkopen.2020.5842
- US Preventive Services Task Force, Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, et al. Screening for lung cancer: US preventive services task force

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1132514/full#supplementary-material>

- recommendation statement. *JAMA* (2021) 325(10):962–70. doi: 10.1001/jama.2021.1117
- Lyu Z, Li N, Chen S, Wang G, Tan F, Feng X, et al. Risk prediction model for lung cancer incorporating metabolic markers: Development and internal validation in a Chinese population. *Cancer Med* (2020) 9(11):3983–94. doi: 10.1002/cam4.3025
- Tu Y, Wu Y, Lu Y, Bi X, Chen T. Development of risk prediction models for lung cancer based on tumor markers and radiological signs. *J Clin Lab Anal* (2021) 35(3):e23682. doi: 10.1002/jcla.23682
- Meza R, Jeon J, Toumazis I, Ten Haaf K, Cao P, Bastani M, et al. Evaluation of the benefits and harms of lung cancer screening with low-dose computed tomography: Modeling study for the US preventive services task force. *JAMA* (2021) 325(10):988–97. doi: 10.1001/jama.2021.1077
- Cai X, Chen L, Kang T, Tang Y, Lim T, Xu M, et al. A prediction model with a combination of variables for diagnosis of lung cancer. *Med Sci Monit Int Med* (2017) 23:5620–9. doi: 10.12659/MSM.904738
- Bae JM. Serum folate levels and lung cancer risk: A meta-epidemiological study of population-based case-control studies. *Asian Pacific J Cancer Prev APJCP* (2020) 21(6):1829–33. doi: 10.31557/APJCP.2020.21.6.1829
- Atasoy O, Cini N, Erdogan MA, Yaprak G, Erbas O. Radiotherapy and high bilirubin may be metformin like effect on lung cancer via possible AMPK pathway modulation. *Bratisl Lek Listy* (2022) 123(2):100–9. doi: 10.4149/BLL_2022_016
- Jin J, Yang L, Liu D, Li WM. Prognostic value of pretreatment lymphocyte-to-Monocyte ratio in lung cancer: A systematic review and meta-analysis. *Technol Cancer Res Treat* (2021) 20:1533033820983085. doi: 10.1177/1533033820983085
- Soeroso NN, Rizki Ananda F, Samosir G, Hariman H, Chairani E, Yanoer P. The correlation between hemostatic parameters and mortality rate in patients with non-small cell lung cancer. *Hematol Rep* (2021) 13(3):8361. doi: 10.4081/hr.2021.8361
- Hammouda A, Souilah S, Ferhat-Hamida MY, Amir ZC, Bouguerra SA, Hariti G, et al. Platelet activation in lung cancer. *Ann Biol Clin (Paris)* (2021) 79(1):41–8. doi: 10.1684/abc.2021.1623



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

Daoming Zhu,
Southern Medical University, China
Xiaoli Liu,
Beijing Ditan Hospital, Capital Medical
University, China

*CORRESPONDENCE

Peng Zeng
✉ zpeng701@163.com

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 12 January 2023

ACCEPTED 28 March 2023

PUBLISHED 14 April 2023

CITATION

Zeng P, Shen D, Shu W, Min S, Shu M,
Yao X, Wang Y and Chen R (2023)
Identification of a novel peptide targeting
TIGIT to evaluate immunomodulation of
 ^{125}I seed brachytherapy in HCC by
near-infrared fluorescence.
Front. Oncol. 13:1143266.
doi: 10.3389/fonc.2023.1143266

COPYRIGHT

© 2023 Zeng, Shen, Shu, Min, Shu, Yao,
Wang and Chen. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Identification of a novel peptide targeting TIGIT to evaluate immunomodulation of ^{125}I seed brachytherapy in HCC by near-infrared fluorescence

Peng Zeng^{1*}, Duo Shen², Wenbin Shu³, Shudan Min⁴, Min Shu⁴,
Xijuan Yao⁵, Yong Wang⁵ and Rong Chen¹

¹Department of Oncology, Zhongda Hospital, Medical School, Southeast University, Nanjing, Jiangsu, China, ²Department of Gastroenterology, The Second People's Hospital of Changzhou, Nanjing Medical University, Changzhou, Jiangsu, China, ³Department of Gastrointestinal Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, China, ⁴Jiangsu Key Laboratory of Molecular and Functional Imaging, Department of Radiology, Zhongda Hospital, Medical School, Southeast University, Nanjing, Jiangsu, China, ⁵Center of Interventional Radiology and Vascular Surgery, Department of Radiology, Zhongda Hospital, Medical School, Southeast University, Nanjing, Jiangsu, China

Introduction: Hepatocellular carcinoma (HCC) has very poor prognosis due to its immunosuppressive properties. An effective measure to regulate tumor immunity is brachytherapy, which uses ^{125}I seeds planted into tumor. T cell immune receptors with immunoglobulin and ITIM domains (TIGIT) is highly expressed in HCC. The TIGIT-targeted probe is expected to be an effective tool for indicating immunomodulation of ^{125}I seed brachytherapy in HCC. In this study, We constructed a novel peptide targeting TIGIT to evaluate the immune regulation of ^{125}I seed brachytherapy for HCC by near-infrared fluorescence (NIRF).

Methods: Expression of TIGIT by immunofluorescence (IF) and flow cytometry (FCM) in different part and different differentiated human liver cancer tissues was verified. An optical fluorescence probe (Po-12) containing a NIRF dye and TIGIT peptide was synthesized for evaluating the modulatory effect of ^{125}I seed brachytherapy. Lymphocytes uptake by Po-12 were detected by FCM and confocal microscopy. The distribution and accumulation of Po-12 in vivo were explored by NIRF imaging in subcutaneous and orthotopic tumors. IHC and IF staining were used to verify the expression of TIGIT in the tumors.

Results: TIGIT was highly expressed in HCC and increased with tumor differentiation. The dye-labeled peptide (Po-12) retained a stable binding affinity for the TIGIT protein *in vitro*. Accumulation of fluorescence intensity (FI) increased with time extended in subcutaneous H22 tumors, and the optimal point is 1 h. TIGIT was highly expressed on lymphocytes infiltrated in tumors and could be suppressed by ^{125}I seed brachytherapy. Accumulation of Po-12-Cy5 was increased in tumor-bearing groups while declined in ^{125}I radiation group.

KEYWORDS

hepatocellular carcinoma, TIGIT, ^{125}I seed brachytherapy, peptide, NIRF

1 Introduction

Hepatocellular carcinoma (HCC) is the fifth most common cancer and the second leading cause of cancer-related death worldwide (1). The treatment of advanced HCC has been a dilemma because of its self-immune tolerance (2). Local radiotherapy (RT) is an effective immunomodulatory measure for tumors (3). ^{125}I seed implantation brachytherapy is a new type of RT that has been widely used in the treatment of a variety of tumors, including liver cancer (4–8). ^{125}I seed is a kind of single miniature radioactive source with low dose rate. The core of this seed is palladium wire of ^{125}I radioactive nuclide, encased in cylindrical sealed titanium alloy tube, with half-life of 59.43 days, average energy of 27 ~ 35 keV and radiation distance of 1.7 ~ 2.0 cm. ^{125}I seed brachytherapy has a good safety profile for the treatment of HCC (9). Increasing evidence has confirmed that ^{125}I seed brachytherapy inhibits tumor growth and activates antitumor immunity (10–13). However, RT alone is not enough to prevent tumor recurrence and metastasis. Combined RT can further promote this immunomodulatory effect, in which the combination of nano-materials such as photothermal therapy (PTT) and photodynamics therapy (PDT) have a significant effect (14–16). More important, RT can also result in immunosuppression with the accumulation of radiation dose (17). Nevertheless, there is currently no accurate method for evaluating immune molecule changes in the tumor microenvironment (TME) for clinical treatment. Therefore, real-time and dynamic monitoring of these molecules is needed to detect changes in immune translation and provide guidance for immunotherapy.

T cell immune receptor with immunoglobulin and ITIM domains (TIGIT) is a receptor of the Ig superfamily. It plays a key role in limiting adaptive and innate immunity and is involved in tumor immune surveillance, mainly expressed on T cells, natural killer cells (NK), and other antigen-presenting cells (APCs), which can reduce cytokine production and show strong immunosuppressive effects (18). Considering that NK cells account for a large proportion in liver, and TIGIT is expressed on both NK and T cells, TIGIT has been reported to be an important inhibitory immune checkpoint (ICP) in HCC (19). RT regulates the expression of TIGIT in tumors. RT combined with anti-TIGIT is a good anti-tumor strategy (20, 21). Consequently, TIGIT can be used as an indicator of tumor immunoactivity. Hence, we used TIGIT as a marker to reflect the immunoregulation of ^{125}I seed brachytherapy in HCC.

The development of molecular imaging technology provides the possibility for the dynamic assessment of tumor immune microenvironment (TIME) changes in real time. It is reasonable to measure immunoactivities by molecular imaging of predictive biomarkers in tumors. Near-infrared fluorescence (NIRF) emitters have been widely used in the real-time imaging of tumors because of their excellent tissue penetration and target-background contrast (22, 23). Therapeutic targeted molecules and immune checkpoints (ICPs) labeled with NIRF dye have been tested in the evaluation of cancer therapy, demonstrating ideal safety and high accuracy in identifying the TME (24–26).

NIR imaging probes are usually composed of NIR dyes and targeting groups (including antibodies and their fragments,

peptides, small molecules, etc.), which can bind to specific molecules in the process of tumorigenesis and development to achieve dynamic tracing of the TME (27–29). Among these, peptides stand out among many targeting groups because of their low immunogenicity, strong tissue penetration, fast blood clearance, and relatively simple production process (30, 31). Phage display technology combines the antigen recognition ability of recombinant proteins and is an efficient screening system to generate peptides against specific molecules or tumor structures. Therefore, it has great prospects in the development of tumor-specific peptides (32).

As TIGIT is highly expressed in HCC, and is a new immunotherapeutic target that may be regulated by RT (33), we introduced ^{125}I seed implantation into the tumor for brachytherapy and regulated the expression of TIGIT. We also designed a 12-amino acid peptide targeting TIGIT to bind to lymphocytes in HCC. This peptide was combined with Cy5 to further evaluate the targeting efficacy of the probe in HCC before and after radiotherapy under NIRF, which can indicate the degree of immune regulation of HCC by ^{125}I seed brachytherapy (Figure 1).

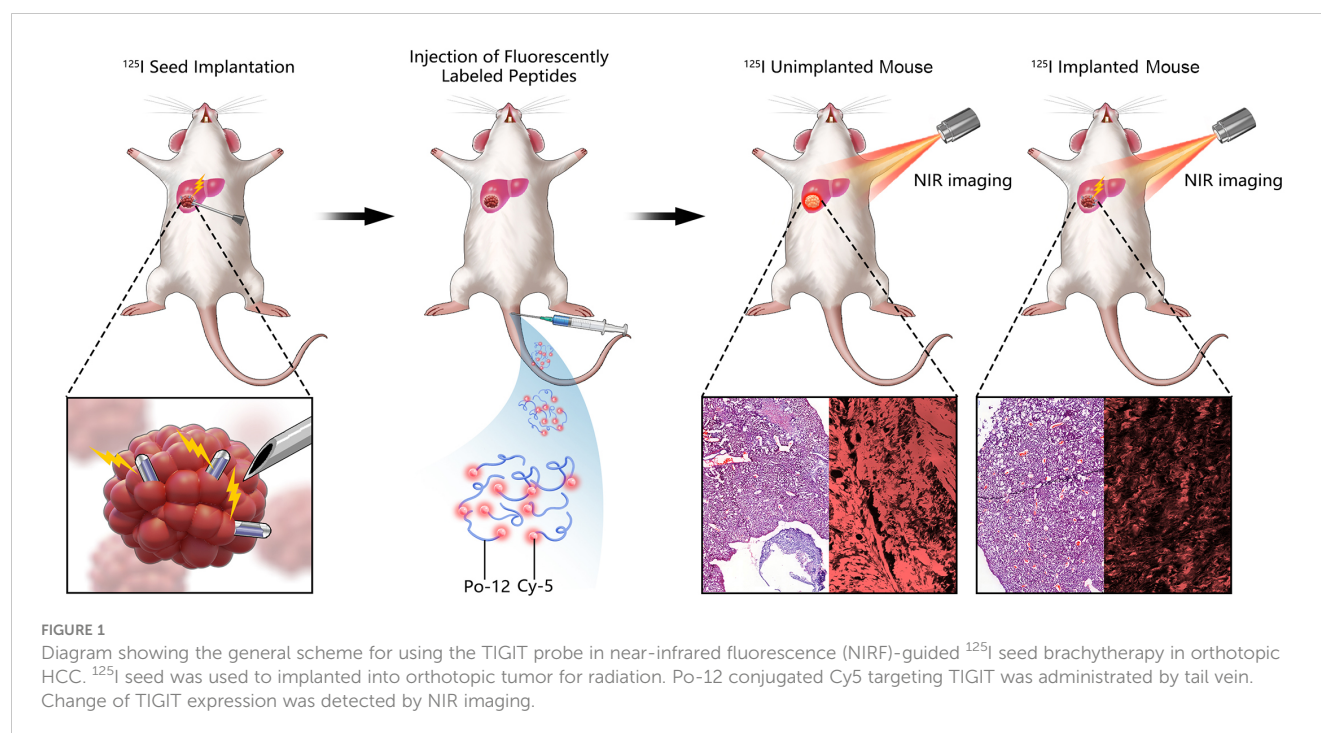
2 Methods

2.1 Expression and purification of recombinant TIGIT antigen

In order to obtain the recombinant TIGIT protein, the TIGIT gene (available in PubMed) was cloned into the BamHI and EcoRI sites of PET28-A vector (+) and transformed into BL21(DE3) competent *Escherichia coli* cells, which were cultured in Luria broth at 37 °C, containing ampicillin (OD values of 0.6–0.8). Subsequently, 1 μM isopropyl-β-D-thio-galactoside (IPTG) was added to the culture to induce protein expression. Bacterial cultures were harvested and centrifuged at 5000 rpm for 10 min, after which the precipitate was resuspended in lysis buffer containing 8 M urea and 50 mM Tris (pH 7.4). After complete decomposition at high pressure, lysed bacteria were centrifuged at 15000 rpm for 30 min and loaded onto a nickel resin-bound (column affinity chromatography) column. The recombinant TIGIT protein was eluted with a highly stringent buffer containing 300 mM imidazole and verified by 10% SDS-PAGE and Coomassie Brilliant Blue staining to obtain purified recombinant TIGIT protein.

2.2 Screening of TIGIT targeted peptides

The resulting purified TIGIT protein was coated onto a 96-well plate and used for subsequent peptide screening. Phage display technology was used for the screening. Fifty microliters of phage supernatant were added to a 96-well plate coated with TIGIT protein for screening. 0.2 M Gly-HCl pH2.2 was used for elution and 1M Tris-HCl pH9.1 was used for neutralization. The neutralization solution (containing bacteriophages) was diluted 1000 or 10000 times and then added to *E. coli* for amplification. After three rounds of screening, the desired affinity clone target was obtained. The cloned target was sequenced using DNA



and the target amino acid sequence was obtained by reverse sequencing. A sequence with a high affinity and the highest occurrence times was selected for chemical synthesis. To better link the Cy5 fluorophore, the Cy5 fluorophore was first added to the synthesized amino acid sequence and then the Cy5 fluorophore was added to the N-terminus of the amino acid sequence. All peptides were chemically synthesized using the solid-phase Fmoc method and purified by high-performance liquid chromatography (HPLC) and electrospray ionization mass spectrometry to a minimum purity of 95%.

2.3 Cell culture and animal models

H22 cells (Chinese National Collection of Authenticated Cell Cultures) were cultured in Roswell Park Memorial Institute 1640 medium (RPMI 1640, Gibco, USA), containing 10% fetal bovine serum (FBS, Gibco, USA) and 10% penicillin/streptomycin (P/S, Gibco, USA) at 37°C in a humidified atmosphere containing 5% CO_2 . All animal experiments were approved by the Animal Ethics Committee of Southeast University and conducted in compliance with the Regulations for the Administration of Affairs Concerning Experimental Animals of China. Six-week-old male BALB/c mice (Vital River Laboratory Animal Technology, China) were housed at the Animal Center of the Southeast University laboratory. 1×10^6 H22 cells were used to induce subcutaneous tumors by an injection into the back of each mouse and tumor tissue was used to generate orthotopic hepatic tumors by implantation into the liver. The mice were anesthetized with an intraperitoneal injection of 60 mg/kg sodium pentobarbital. ^{125}I seeds (activity of 0.8 mCi) were implanted into the tumor for radiation.

2.4 Cell sorting of T lymphocytes

The tumor tissues of each group were minced into small fragments and digested with tissue digestive enzymes (Miltenyi Biotec, Germany) at 37°C for 40 min. Single cells were collected by filtration through a $70\text{ }\mu\text{m}$ colander (BD Biosciences, USA). T lymphocytes were sorted using magnetic beads and a $\text{CD}45^{+}$ lymphocyte isolation kit (STEMCELL Technologies, Canada).

2.5 Flow cytometry

Isolated $\text{CD}45^{+}$ lymphocytes were collected at 1×10^6 cells/sample and incubated with 300 μL control Con-12 or Po-12 (10 $\mu\text{g}/\text{mL}$) at 4°C for 15 min. After incubation, the cells were washed with PBS 3 times and then resuspended in 400 μL staining buffer. Fluorescence analysis was performed using a flow cytometer (BD Biosciences) with a count of 1×10^6 living cells per sample. The results were analyzed by flow cytometry using Flow Jo software for 3 times (v7.6, OR, USA).

2.6 Immunofluorescence staining

Isolated $\text{CD}45^{+}$ lymphocytes were seeded in a confocal chamber at 1×10^6 cells/well for 24 h and fixed with 4% paraformaldehyde at room temperature for 20 min. The cells were incubated with Po-12-Cy5 or Con-12-Cy5 at 4°C overnight. After staining with 4' 6-diamidino-2-phenylindole (DAPI), cells were imaged using a confocal microscope (FV3000; Olympus, Japan). The prepared

tumor sections were also subjected to immunofluorescence (IF), and the tumor tissues were resected and frozen for IF imaging. Slides were stained with DAPI and analyzed using a confocal microscope (FV3000, Olympus, Japan).

2.7 Western blotting

Isolated CD45⁺ lymphocytes were lysed to concentrate proteins using RIPA lysis buffer (Beyotime Biotechnology, China). Cell extracts were clarified by centrifugation, and protein concentrations were determined using the BCA assay. Protein extracts were separated by SDS-PAGE, transferred to microporous polyvinylidene difluoride membranes (Roche, USA), and blocked using 5% BSA. Then the membranes were incubated with anti-TIGIT polyclonal antibody (Abexxa, UK) or GAPDH monoclonal antibody (Cell Signaling Technology, MA, USA) at 4°C overnight. After washing, the membranes were incubated with HRP-conjugated secondary antibodies (Cell Signaling Technology, USA) at room temperature for 1 h. Protein bands were detected with enhanced chemiluminescence (ECL) and imaged using a chemiluminescence system (Bio-Rad, USA). The above experimental procedures were repeated 3 times.

2.8 Near-infrared fluorescence imaging

Six H22 tumor-bearing mice were randomly divided into two groups and intravenously injected with 20 µg Po-12-Cy5-peptide or Con-12-Cy5-peptide. After anesthesia with isoflurane in oxygen, *in vivo* fluorescence imaging was performed using an IVIS-Spectrum system (Perkin Elmer, Santa Clara, CA, USA) at several time points (0.5, 0.75, 1, 2, 4, and 8 h). The excitation and emission wavelengths of the probe were 620 and 670 nm, respectively. The mice were sacrificed after injection of the peptide; their tumors and major organs were dissected for *ex vivo* NIR imaging.

2.9 Statistical analysis

All data are presented as mean ± standard deviation (SD). Statistical significance between groups was determined using two-tailed Student's t-test or one-way analysis of variance (ANOVA). The threshold of statistical significance was set at $P < 0.05$ (* $P < 0.05$, ** $P < 0.01$). Statistical analyses were performed using the GraphPad Prism software (V9.0, CA, USA).

3 Results

3.1 TIGIT expression in human HCC samples

In experiments examining the expression of TIGIT protein in human HCC, different parts of the tissue (normal, paracancer, and

tumor tissue) and differentiated tumor tissues (well, moderate, and poorly differentiated tumors) were collected. H&E and IF staining was used to assess TIGIT expression in each group. The probe distribution assay of IF showed extensive accumulation of fluorescence in tumor tissue compared to paracancerous tissue but was negligible in normal tissue (Figure 2A). Furthermore, fluorescence intensity (FI) increased with the degree of malignancy, which revealed a stronger accumulation in poorly differentiated tumors than in moderately differentiated tumors, while the well-differentiated tumor displayed the least FI (Figure 2C). To further verify the expression of TIGIT, we detected the expression of TIGIT on the surface of lymphocytes in each tissue using flow cytometry (FCM), and the results showed a consistent trend (Figures 2B–D, $P < 0.01$). These results indicate that TIGIT was highly expressed in HCC and increased with tumor differentiation.

3.2 Identification and synthesis of TIGIT-targeted peptides

TIGIT protein was successfully purified. Three rounds of incubation and screening of expressed proteins were performed (Figure 3A). Significant enrichment of recovered phages was observed (Figure 3B). A consistently predicted molecular weight of approximately 29 kDa was determined by Coomassie Brilliant Blue staining (Figure 3C). After the last round of screening, 19 clones were randomly selected, verified by enzyme-linked immunosorbent assay (ELISA), and sequenced by high-throughput sequencing. The absorbance of the two highly repetitive peptide sequences at 450 nm was significantly higher than that of the control sample (Figure 3D). The high frequency of the peptide sequences indicated efficient enrichment during the screening process. The peptide sequence GAQYPHISRALH (named Po-12), with an OD equal to 10 times that of the control, was selected as the best candidate peptide for subsequent studies, and the peptide with sequence shuffling (named Con-12) was used as the control peptide (Figure 2D). The molecular structure of Po-12-Cy5 is presented in Figure S1A. Cy5 fluorophore was added to (red marker) the N-terminus of the naked peptide (Figure S1A). The mass to charge ratio (M/Z) of Po-12-Cy5 by mass spectrometry was determined at 2231.58 (Figure S1B), and the retention time of peptide purification by HPLC was 11.023 min (Figure S1C).

3.3 Binding of TIGIT-targeted peptide to lymphocytes

The *in vitro* specificity of Po-12 to the TIGIT protein was evaluated using FCM. The results revealed that lymphocytes in the Po-12 group showed a stronger absorption of fluorescence intensity than those in the Con-12 group and isotype group (Figure 4A, $P < 0.01$). Confocal microscopy imaging was used to evaluate the cellular binding of the probes. Strong membranous binding was observed in lymphocytes treated with Po-12-Cy5, whereas almost no fluorescence was found in Con-12-Cy5-treated one (Figure 4B).

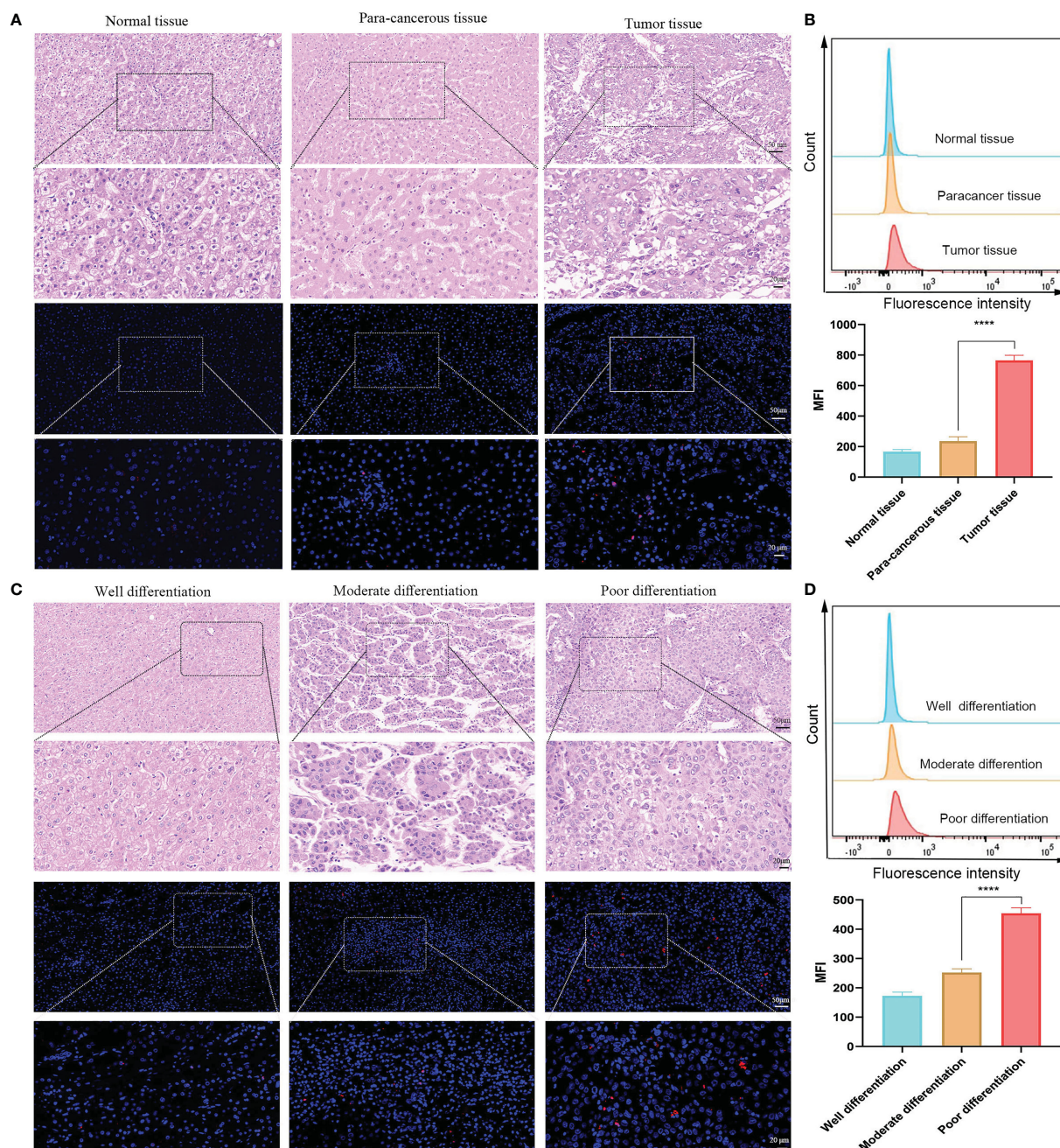


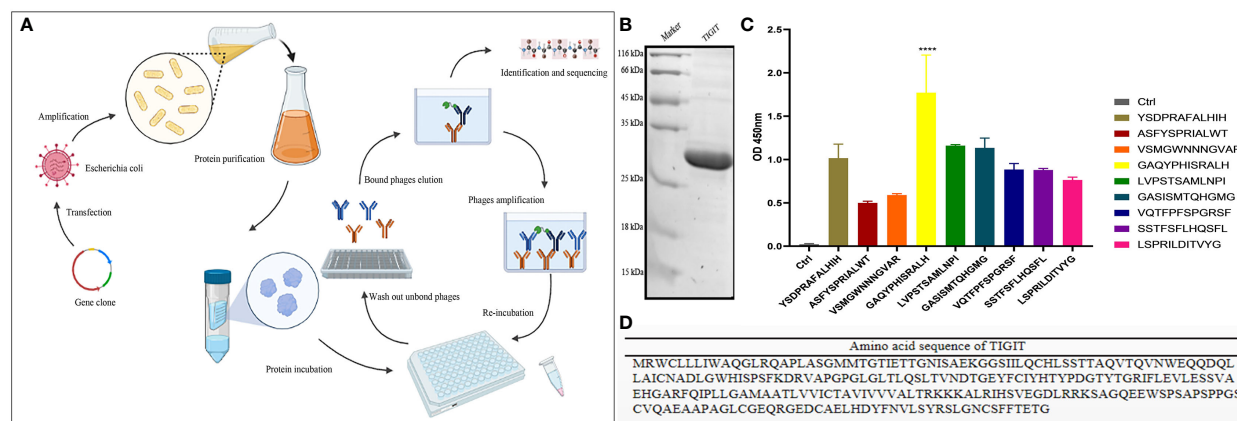
FIGURE 2
TIGIT expression in human HCC. (A) Representative IF staining of TIGIT in different parts of tissue. (B) Flow cytometric analysis of TIGIT expression on lymphocytes in different parts of tissue (n=3). (C) Representative IF staining of TIGIT in various differentiated tumor tissue. (D) Flow cytometric analysis of TIGIT expression on lymphocytes in different differentiated tumor tissue (n=3). (Scale bar: Up, 50 μ m; Down, 20 μ m). ****P < 0.01.

These data indicated that the dye-labeled peptide retained a stable binding affinity for the TIGIT protein *in vitro*.

3.4 NIRF imaging of tumors models

NIRF imaging was performed in tumor-bearing BALB/c mice by intravenous injection of Cy5-peptides. Accumulation of FI increased in subcutaneous H22 tumors from 0.5 h until 1 h, after

which the FI began declining. Quantitative analysis showed that the mean fluorescence intensity (MFI) of Po-12-Cy5 was significantly higher than that of Con-12-Cy5 (Figures 5A, B, P < 0.01). Ex vivo optical imaging of the tumors and main organs was performed 1 h post-injection. The quantification of FI corroborated the visualization of *in vivo* optical imaging. Biodistribution analysis indicated that Po-12-Cy5 showed prominent renal clearance. The enrichment of Po-12-Cy5 in H22 tumors was the highest in all organs except the heart (Figures 5C, D).



3.5 TIGIT expression and cellular uptake of probes in lymphocytes after brachytherapy

To detect the effect of brachytherapy on the expression of TIGIT in HCC, we established mouse subcutaneous tumor models under ^{125}I seed radiation. Lymphocytes from each group were isolated from tumor tissue using magnetic beads and incubated with Po12-Cy5. Confocal microscopy imaging was also used to evaluate the cellular uptake of the TIGIT-targeted probe in lymphocytes after radiation. The results showed that the expression of TIGIT in the tumor control group was significantly increased compared with that in the normal control group and decreased in the ^{125}I seed radiation group (Figure 6A). Consistent with confocal microscopy images, TIGIT protein expression in each group extracted from isolated lymphocytes assessed by western blotting also exhibited a similar tendency (Figures 6B, C, $P < 0.01$). These results demonstrated that TIGIT was highly expressed on lymphocytes infiltrated in tumors and could be suppressed by ^{125}I seed brachytherapy, which suggested that this probe may not only visualize the expression changes of TIGIT in tumors but also provide dynamic guidance for RT in TME regulation.

3.6 NIRF-guided TIGIT expression in tumor model of HCC after brachytherapy

To verify the effect of radiation on TIGIT expression and the targeting of the probe in HCC, subcutaneous and orthotopic HCC tumor models were established in mice. The probe was injected into the tail vein, and accumulation of Po-12-Cy5 was increased in both subcutaneous and orthotopic tumors, while the normal control group without tumor showed no change. Quantification analysis revealed that FI in ^{125}I radiation group declined significantly than tumor control one (Figures 7A–G, $P < 0.01$). The results of tumor growth curve were also consistent with IF, and the growth of

subcutaneous and orthotopic tumor was inhibited observably (Figures 7C–H, $P < 0.01$). In addition, the tumors were excised for IHC and IF staining. Results revealed that TIGIT protein expression in IHC was negligible in the normal control group, while it was highly expressed in the tumor-bearing group and downregulated in the ^{125}I radiation group (Figures 7D–I). In agreement with IHC, extensive accumulation of Po-12-Cy5 was observed in tumor-bearing groups compared to that in the normal control group, and also declined in the ^{125}I radiation group (Figures 7E–J). These results reveal that Po-12-Cy5 has excellent TIGIT-positive tumor-targeting potential and can be used as a significant indicator of radiation immune regulation.

4 Discussion

Evaluation of the immune response has always been a challenge in tumor therapy because of the potential reversion or pseudoprogression (34). However, there is still a lack of effective non-invasive real-time dynamic evaluation measures based on immune molecules. The advent of molecular imaging technology provides an opportunity for noninvasive observation of abnormal immune molecular events *in vivo*. In response to the expression of TIGIT protein in cancer, a variety of therapeutic antibodies have been developed in phase I-II clinical trials (35, 36). Probe navigation systems for recognizing tumor molecules mainly include antibodies, peptides, and small molecules (27–29). Among them, peptides have been valued for their ability to bind hidden epitopes because of their smaller molecular weight (31). With the emergence and rapid development of phage display technology, new peptidyl molecular probes have greatly promoted the detection of tumor molecules, showing great potential for clinical exploration (37, 38).

To the best of our knowledge, the indication of ICP by optical labeling of peptide targeting TIGIT under real-time NIF for RT regulation has not been reported. In this study, the target peptide of

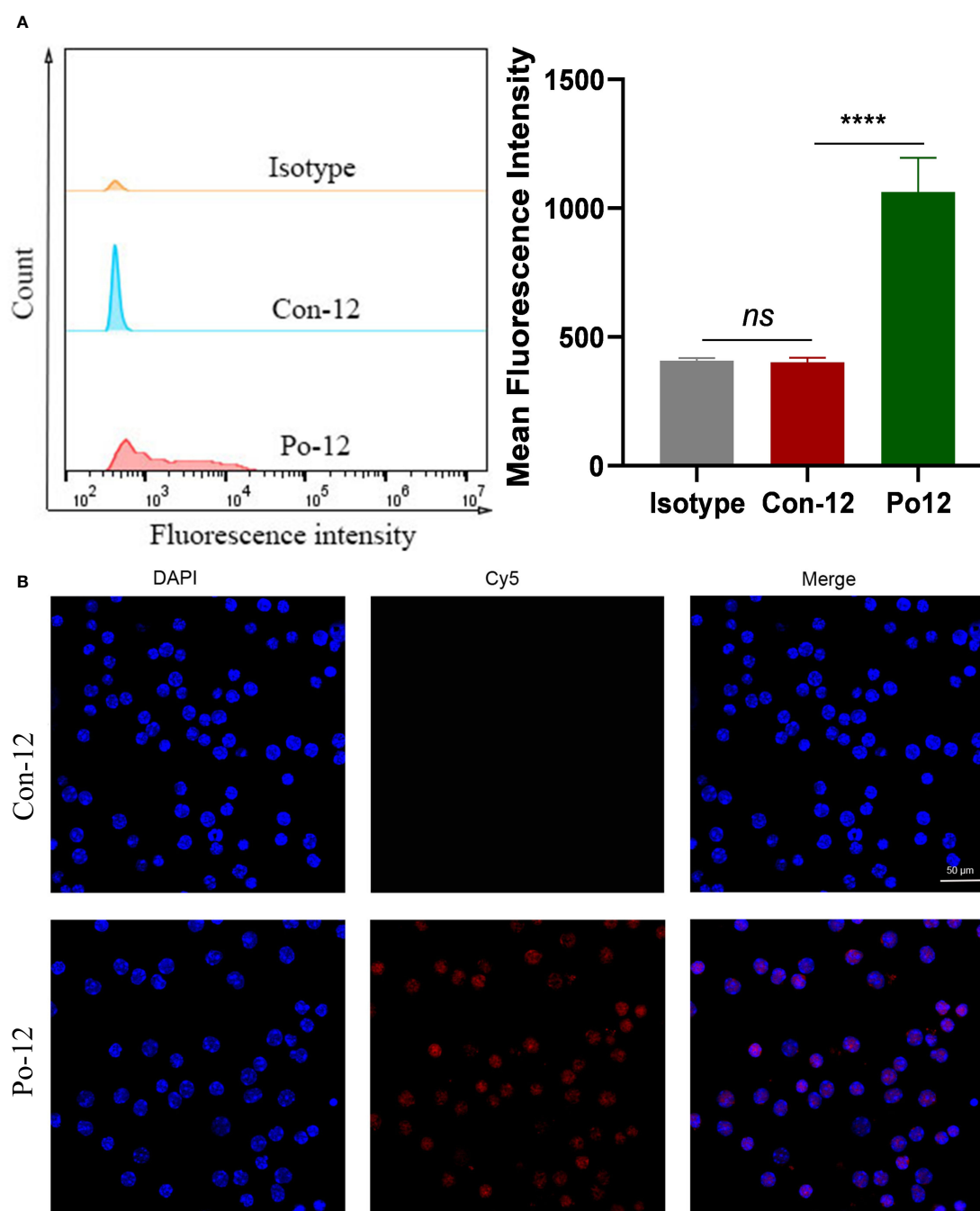


FIGURE 4
Binding of TIGIT-targeted peptide to lymphocytes. **(A)** FCM analysis of lymphocytes after incubation with Con-12-Cy5 or Po-12-Cy5 peptide ($n=3$). **(B)** Confocal images of lymphocytes after treatment with Con-12-Cy5 or Po-12-Cy5 peptide. Scale bar: 50 μ m. ns, no significance; **** $P < 0.01$.

TIGIT was identified using phage display technology. To determine the expression of TIGIT in HCC, we first collected different parts and differentiated human HCC tissues and detected the expression of TIGIT by IF staining and FCM. The results showed that the expression of TIGIT in tumor tissues was significantly higher than that in adjacent and normal tissues, and the expression of TIGIT in poorly differentiated liver cancer tissues was significantly higher than that in moderately and well-differentiated tissues. Furthermore, we isolated lymphocytes from mouse tumors and

verified the high binding affinity of the Cy5-conjugated peptide (Po-12-Cy5) to lymphocytes by FCM and immunofluorescence confocal assay *in vitro*. *In vivo* optical imaging further verified the targeting ability of Po-12-Cy5 in a subcutaneous HCC model. The results showed that the fluorescence uptake of Po-12-Cy5 was significantly stronger than that of Con-12-Cy5, peaking rapidly within 1 h, and gradually declining over 7 h. In addition, the biodistribution results showed that the fluorescence intensity in the tumor was significantly higher than that of the hybrid peptide, which was

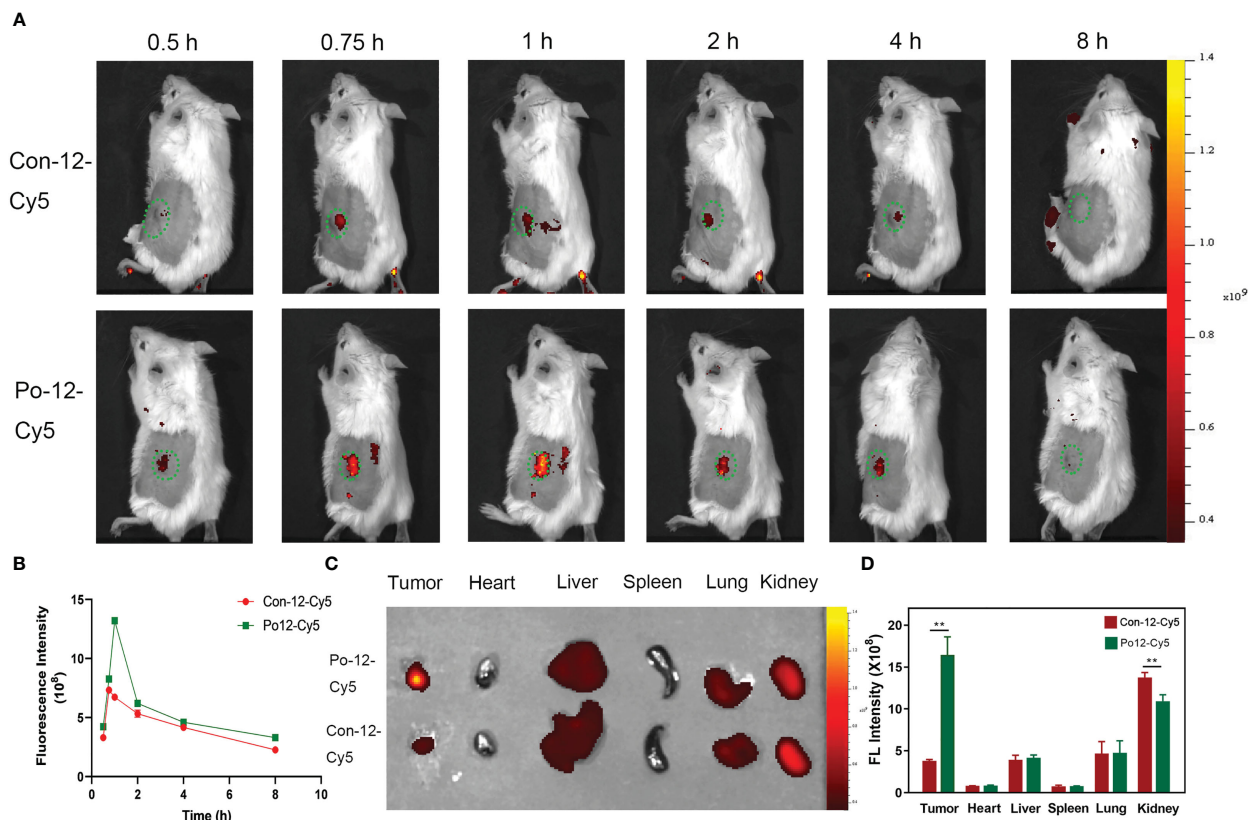


FIGURE 5

In vivo imaging of H22 subcutaneous tumors and biodistribution of the probe. (A) *In vivo* imaging post-injection of probes and (B) quantification of fluorescence intensity ($n=3$). (C) *Ex vivo* imaging of tumor and normal organs (Tumor, heart, liver, spleen, lung, kidney) and (D) quantification of fluorescence intensity ($n=3$). ** $P < 0.01$.

consistent with *in vivo* observations. In addition, the uptake of Po-12-Cy5 in the kidney was higher than that in other organs, indicating that the kidney may be the main excretion route. This is consistent with what has been reported in a series of other literatures (39–41). These results indicate that Po-12-Cy5 has a good targeting effect on TIGIT in HCC and can be used as an important indicator of changes in the immune microenvironment in HCC. Therefore, we planned to use it to dynamically monitor the regulation of TIGIT by ^{125}I seed RT in real-time.

Based on the specificity of the Po-12-Cy5 probe for TIGIT in liver cancer, we constructed a tumor brachytherapy model using ^{125}I seed implantation and isolated tumor-infiltrating lymphocytes and compared the expression of TIGIT under ^{125}I seed RT, tumor control, and normal control groups *in vitro*. Immunofluorescence confocal analysis showed that the fluorescence uptake in the tumor group was significantly higher than that in the normal control group, while that in the ^{125}I seed RT group was downregulated. The expression of the TIGIT protein in each group also demonstrated this trend. To further demonstrate the indication of the Po-12-Cy5 probe on TIGIT by RT, we constructed subcutaneous and orthotopic tumor models of HCC in mice and further evaluated the targeting of the Po-12-Cy5 probe *in vivo*. Quantitative analysis showed that the FI of the ^{125}I seed RT group was significantly lower than that of the tumor control group in both the subcutaneous and

orthotopic tumor models. IF and IHC staining of tumor tissues also showed this trend. These results indicate that the Po-12-Cy5 probe has precise targeting of TIGIT in HCC and can be used as an indicator of RT immunoregulation, which has important clinical significance for guiding HCC immunotherapy.

However, there are still some problems associated with the clinical translation of the Po-12-Cy5 probe. First, the clinical safety of the NIF dye-Cy5, was not confirmed. Nonetheless, NIR-II dye-indocyanine green (ICG) has been approved for clinical use by the Food and Drug Administration (FDA) (42, 43). IRDye800cw, as a marker of SHRMab antibodies, has been widely used because it has no obvious clinical toxicity evaluated in human trials (44). Therefore, it provides the possibility of improving the clinical translation of the Po-12 probe. Second, although peptides show superior performance in tumor diagnostic applications, their binding affinity is not yet comparable to that of specific antibodies. At present, the antibody-drug conjugate (ADC) has been used in tumor therapy as a very promising antitumor drug because of its high affinity and targeting (45). Therefore, using TIGIT as a naked antibody of ADC and further conjugation of NIR-II dye with peptide can not only further solve the limitations of this study but also further improve the efficacy of HCC immunotherapy. Finally, in contrast to bioluminescence imaging, which is affected by tissue depth and imaging dimension, PET/SPE-CT technology for

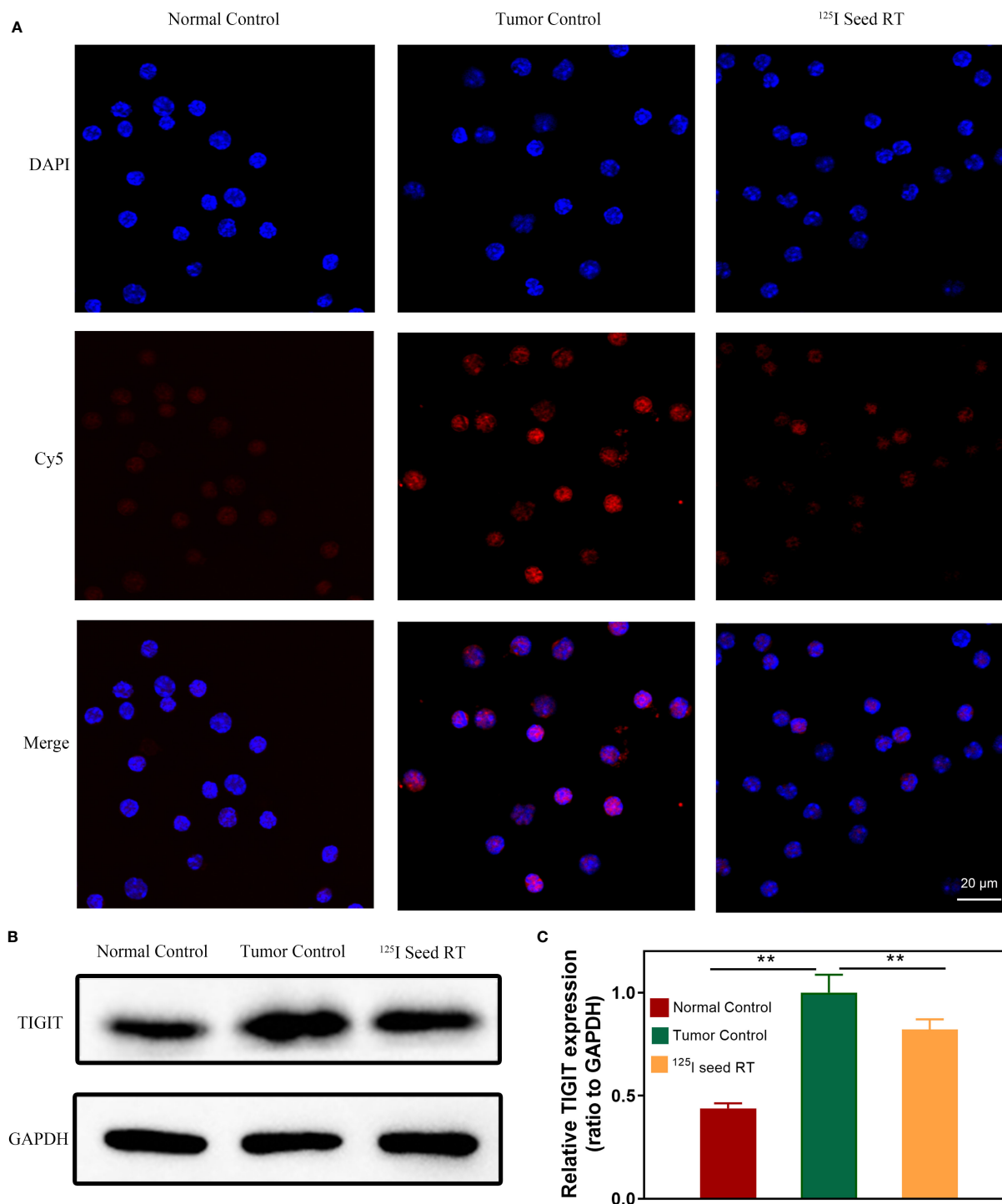


FIGURE 6
TIGIT expression and cellular uptake of probes in lymphocytes. **(A)** Confocal microscopic imaging of the cellular binding of probes in lymphocytes. **(B)** Western blotting of TIGIT protein in lymphocytes and **(C)** quantification of TIGIT expression ($n=3$). Scale bar: 20 μm . ** $P < 0.01$.

small animals can achieve absolute quantification owing to the excellent penetration ability of radionuclides, with no signal attenuation; this provides three-dimensional information and accurate localization (46). Therefore, we need to construct PET imaging probes to evaluate tumor immune molecules in future studies.

5 Conclusion

In this study, we synthesized a TIGIT targeting NIRF probe, Po-12-Cy5. *In vitro* and *in vivo* experiments showed that Po-12-Cy5 was specifically absorbed by infiltrating lymphocytes in HCC. In addition, the probe could indicate TIGIT regulation of ^{125}I seed radiation under NIRF

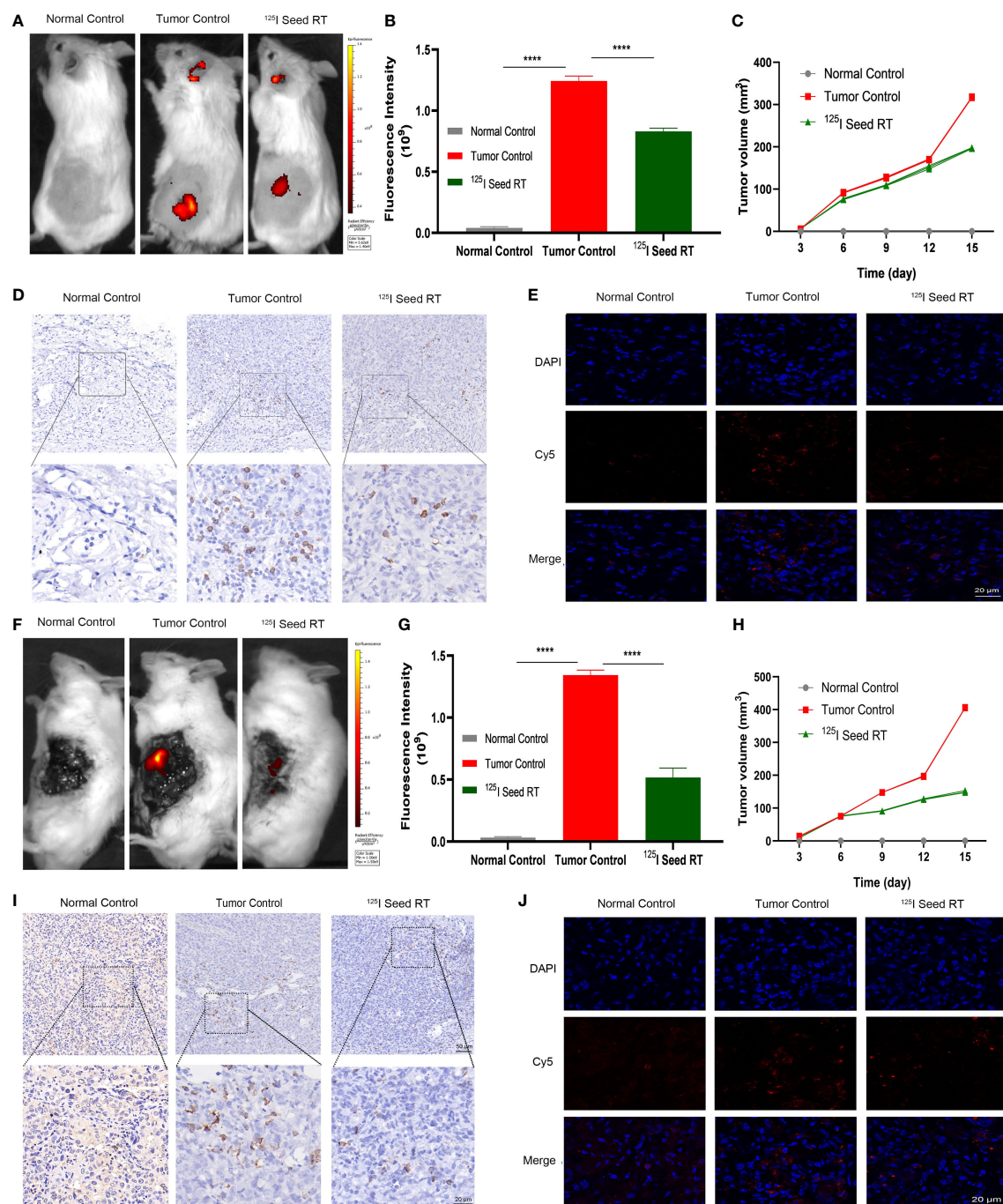


FIGURE 7

Fluorescence imaging in subcutaneous and orthotopic model of HCC. (A–F) Imaging of subcutaneous and orthotopic tumor after infusion of probes. (B–G) Quantification of FI in subcutaneous and orthotopic tumor ($n=3$). (D–I) IHC of TIGIT expression in subcutaneous and orthotopic tumor. (E–J) Probe distribution evaluated via IF staining of frozen sections from subcutaneous and orthotopic tumor. (C–H) Tumor volume of subcutaneous and orthotopic tumor after ^{125}I seed radiation ($n=4$). Scale bar: 50 μm . Scale bar: 20 μm . **** $P < 0.01$.

guidance. As the TIGIT protein correlates with the degree of tumor differentiation and can be downregulated by RT, we believe that this probe can help indicate the regulation of the immune microenvironment of HCC by RT. Therefore, the Po-12-Cy5 probe can be used as an effective immunoevaluation tool with clinical translational potential.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by IEC for Clinical Research of Zhongda Hospital, Affiliated to Southeast University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Animal Experimental Ethical Inspection Form of Southeast University.

Author contributions

PZ, DS, and WS designed the experiments and carried out data analyses. SM, XY, and MS collated the data and performed the statistical analysis. YW and RC produced the initial draft of the manuscript. PZ finalized the manuscript and had final responsibility for the decision to submit for publication. All authors contributed to the article and approved the submitted version.

Funding

The study was supported by Jiangsu Provincial Special Program of Medical Science (BE2019750), National Natural Science Foundation of China (81827805), National Key Research and Development Program (2018YFA0704100, 2018YFA0704104), Science Fund for Creative Research Groups of the National Natural Science Foundation of China (61821002) and

Interventional radiology scientific research special fund project of Jiangsu medical association (SYH-3201140-0025(2021020)) for their financial supports.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1143266/full#supplementary-material>

References

- Vogel A, Meyer T, Sapisochin G, Salem R, Saborowski A. Hepatocellular carcinoma. *Lancet* (2022) 400(10360):1345–62. doi: 10.1016/s0140-6736(22)01200-4
- Jenne CN, Kubes P. Immune surveillance by the liver. *Nat Immunol* (2013) 14(10):996–1006. doi: 10.1038/ni.2691
- Rodriguez-Ruiz ME, Vitale I, Harrington KJ, Melero I, Galluzzi L. Immunological impact of cell death signaling driven by radiation on the tumor microenvironment. *Nat Immunol* (2020) 21(2):120–34. doi: 10.1038/s41590-019-0561-4
- Lu J, Guo JH, Zhu HD, Zhu GY, Chen L, Teng GJ. Safety and efficacy of irradiation stent placement for malignant portal vein thrombus combined with transarterial chemoembolization for hepatocellular carcinoma: A single-center experience. *J Vasc Interv Radiol* (2017) 28(6):786–94.e3. doi: 10.1016/j.jvir.2017.02.014
- Zhu HD, Guo JH, Huang M, Ji JS, Xu H, Lu J, et al. Irradiation stents vs. conventional metal stents for unresectable malignant biliary obstruction: A multicenter trial. *J Hepatol* (2018) 68(5):970–7. doi: 10.1016/j.jhep.2017.12.028
- Guo JH, Teng GJ, Zhu GY, He SC, Fang W, Deng G, et al. Self-expandable esophageal stent loaded with 125I seeds: Initial experience in patients with advanced esophageal cancer. *Radiology* (2008) 247(2):574–81. doi: 10.1148/radiol.2472070999
- Jiao D, Zhou X, Yao Y, Chen J, Lei Q, Ren J, et al. 125I seed implantation brachytherapy for glottic carcinoma: An experimental and clinical study. *Am J Cancer Res* (2021) 11(4):1321–34.
- Kono Y, Kubota K, Mitsumoto T, Tanaka A, Ishibashi A, Kobayashi K, et al. Scintigraphic detection of 125I seeds after permanent brachytherapy for prostate cancer. *J Nucl Med* (2008) 49(4):541–5. doi: 10.2967/jnumed.107.046474
- Li S, Li L, Li B, Wang W. Safety and efficacy of endovascular implantation of a portal vein stent combined with iodine-125 seed-strips followed by transcatheter arterial chemoembolization with sorafenib for the treatment of hepatocellular carcinoma with portal vein tumor thrombosis. *Br J Radiol* (2020) 93(1112):20190279. doi: 10.1259/bjr.20190279
- Zhang FJ, Li CX, Jiao DC, Zhang NH, Wu PH, Duan GF, et al. Ct guided 125iodine seed implantation for portal vein tumor thrombus in primary hepatocellular carcinoma. *Chin Med J* (2008) 121(23):2410–4. doi: 10.1097/00029330-200812010-00009
- Xiang GA, Chen KY, Wang HN, Xiao JF. Immunological influence of iodine-125 implantation in patients with hepatocellular carcinoma resection. *Nan Fang Yi Ke Da Xue Xue Bao* (2010) 30(2):292–4.
- Han T, Yang X, Xu Y, Zheng Z, Yan Y, Wang N. Therapeutic value of 3-d printing template-assisted (125)I-seed implantation in the treatment of malignant liver tumors. *Onco Targets Ther* (2017) 10:3277–83. doi: 10.2147/ott.S134290
- Xie SL, Wu ZY, Zhou L, Liang YQ, Wang XH, Niu LZ, et al. Iodine-125 seed implantation and allogenic natural killer cell immunotherapy for hepatocellular carcinoma after liver transplantation: A case report. *Oncotargets Ther* (2018) 11:7345–52. doi: 10.2147/Ott.S166962
- Huang C, Liu Z, Chen M, Du L, Liu C, Wang S, et al. Tumor-derived biomimetic nanozyme with immune evasion ability for synergistically enhanced low dose radiotherapy. *J Nanobiotechnology* (2021) 19(1):457. doi: 10.1186/s12951-021-01182-y
- Doix B, Trempele N, Riant O, Feron O. Low photosensitizer dose and early radiotherapy enhance antitumor immune response of photodynamic therapy-based dendritic cell vaccination. *Front Oncol* (2019) 9:811. doi: 10.3389/fonc.2019.00811
- Ji C, Zhao M, Wang C, Liu R, Zhu S, Dong X, et al. Biocompatible tantalum nanoparticles as radiosensitizers for enhancing therapy efficacy in primary tumor and metastatic sentinel lymph nodes. *ACS Nano* (2022) 16(6):9428–41. doi: 10.1021/acsnano.2c02314
- Oweida A, Hararah MK, Phan A, Binder D, Bhatia S, Lennon S, et al. Resistance to radiotherapy and pd-L1 blockade is mediated by Tim-3 upregulation and regulatory T-cell infiltration. *Clin Cancer Res Off J Am Assoc Cancer Res* (2018) 24(21):5368–80. doi: 10.1158/1078-0432.Ccr-18-1038
- Freed-Pastor WA, Lambert LJ, Ely ZA, Pattada NB, Bhutkar A, Eng G, et al. The Cd155/Tigit axis promotes and maintains immune evasion in neoantigen-expressing pancreatic cancer. *Cancer Cell* (2021) 39(10):1342–60.e14. doi: 10.1016/j.ccell.2021.07.007

19. Chiu DK, Yuen VW, Cheu JW, Wei LL, Ting V, Fehlings M, et al. Hepatocellular carcinoma cells up-regulate Pvr11, stabilizing pvr and inhibiting the cytotoxic T-cell response *Via* tigit to mediate tumor resistance to Pd1 inhibitors in mice. *Gastroenterology* (2020) 159(2):609–23. doi: 10.1053/j.gastro.2020.03.074
20. Grapin M, Richard C, Limagne E, Boidot R, Morgand V, Bertaut A, et al. Optimized fractionated radiotherapy with anti-Pd-L1 and anti-tigit: A promising new combination. *J Immunother Cancer* (2019) 7(1):160. doi: 10.1186/s40425-019-0634-9
21. Choi C, Yoo GS, Cho WK, Park HC. Optimizing radiotherapy with immune checkpoint blockade in hepatocellular carcinoma. *World J Gastroenterol* (2019) 25(20):2416–29. doi: 10.3748/wjg.v25.i20.2416
22. Lamberts LE, Koch M, de Jong JS, Adams ALL, Glatz J, Kranendonk MEG, et al. Tumor-specific uptake of fluorescent bevacizumab-IrDye800cw microdosing in patients with primary breast cancer: A phase I feasibility study. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2017) 23(11):2730–41. doi: 10.1158/1078-0432.Ccr-16-0437
23. Rosenthal EL, Warram JM, de Boer E, Chung TK, Korb ML, Brandwein-Gensler M, et al. Safety and tumor specificity of cetuximab-IrDye800 for surgical navigation in head and neck cancer. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2015) 21(16):3658–66. doi: 10.1158/1078-0432.Ccr-14-3284
24. Du Y, Sun T, Liang X, Li Y, Jin Z, Xue H, et al. Improved resection and prolonged overall survival with pd-1-IrDye800cw fluorescence probe-guided surgery and pd-1 adjuvant immunotherapy in 4t1 mouse model. *Int J Nanomedicine* (2017) 12:8337–51. doi: 10.2147/ijn.S149235
25. Wan H, Ma H, Zhu S, Wang F, Tian Y, Ma R, et al. Developing a bright nir-ii fluorophore with fast renal excretion and its application in molecular imaging of immune checkpoint PD-L1. *Adv Funct Mater* (2018) 28(50):1804956. doi: 10.1002/adfm.201804956
26. Zhang M, Jiang H, Zhang R, Jiang H, Xu H, Pan W, et al. Near-infrared fluorescence-labeled anti-Pd-L1-Mab for tumor imaging in human colorectal cancer xenografted mice. *J Cell Biochem* (2019) 120(6):10239–47. doi: 10.1002/jcb.28308
27. Li D, Yang M, Liang M, Mei C, Lin Y, Yang F, et al. C-Met-Targeted near-infrared fluorescent probe for real-time depiction and dissection of perineural invasion and lymph node metastasis lesions in pancreatic ductal adenocarcinoma xenograft models. *Biomater Sci* (2021) 9(20):6737–52. doi: 10.1039/d1bm00674f
28. Xiao YT, Zhou C, Ye JC, Yang XC, Li ZJ, Zheng XB, et al. Integrin A6-targeted positron emission tomography imaging of colorectal cancer. *ACS Omega* (2019) 4(13):15560–6. doi: 10.1021/acsomega.9b01920
29. Jiang WL, Wang WX, Wang ZQ, Tan M, Mao GJ, Li Y, et al. A tumor-targeting near-infrared fluorescent probe for real-time imaging atp in cancer cells and mice. *Anal Chim Acta* (2022) 1206:339798. doi: 10.1016/j.aca.2022.339798
30. Zhang P, Cui Y, Anderson CF, Zhang C, Li Y, Wang R, et al. Peptide-based nanoprobe for molecular imaging and disease diagnostics. *Chem Soc Rev* (2018) 47(10):3490–529. doi: 10.1039/c7cs00793k
31. Wang W, Hu Z. Targeting peptide-based probes for molecular imaging and diagnosis. *Adv Mater* (2019) 31(45):e1804827. doi: 10.1002/adma.201804827
32. Saw PE, Song EW. Phage display screening of therapeutic peptide for cancer targeting and therapy. *Protein Cell* (2019) 10(11):787–807. doi: 10.1007/s13238-019-0639-7
33. Solomon BL, Garrido-Laguna I. Tigit: A novel immunotherapy target moving from bench to bedside. *Cancer Immunol Immunother* (2018) 67(11):1659–67. doi: 10.1007/s00262-018-2246-5
34. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekas S, et al. Irecist: Guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* (2017) 18(3):e143–e52. doi: 10.1016/s1470-2045(17)30074-8
35. Cho BC, Abreu DR, Hussein M, Cobo M, Patel AJ, Secen N, et al. Tiragolumab plus atezolizumab versus placebo plus atezolizumab as a first-line treatment for pd-L1-Selected non-Small-Cell lung cancer (Cityscape): Primary and follow-up analyses of a randomised, double-blind, phase 2 study. *Lancet Oncol* (2022) 23(6):781–92. doi: 10.1016/s1470-2045(22)00226-1
36. Niu J, Maurice-Dror C, Lee DH, Kim DW, Nagrial A, Voskoboinik M, et al. First-in-Human phase 1 study of the anti-tigit antibody vibostolimab as monotherapy or with pembrolizumab for advanced solid tumors, including non-Small-Cell lung cancer(☆). *Ann Oncol Off J Eur Soc Med Oncol* (2022) 33(2):169–80. doi: 10.1016/j.annonc.2021.11.002
37. Krumpke LR, Mori T. The use of phage-displayed peptide libraries to develop tumor-targeting drugs. *Int J Pept Res Ther* (2006) 12(1):79–91. doi: 10.1007/s10989-005-9002-3
38. Kumar R, Parray HA, Shrivastava T, Sinha S, Luthra K. Phage display antibody libraries: A robust approach for generation of recombinant human monoclonal antibodies. *Int J Biol Macromol* (2019) 135:907–18. doi: 10.1016/j.jbiomac.2019.06.006
39. Feng GK, Ye JC, Zhang WG, Mei Y, Zhou C, Xiao YT, et al. Integrin A6 targeted positron emission tomography imaging of hepatocellular carcinoma in mouse models. *J Control Release* (2019) 310:11–21. doi: 10.1016/j.jconrel.2019.08.003
40. Li Y, Zhang D, Shi Y, Guo Z, Wu X, Ren JL, et al. Syntheses and preliminary evaluation of [(18) F]Alf-Nota-G-Tmp1 for pet imaging of high aggressive hepatocellular carcinoma. *Contrast Media Mol Imaging* (2016) 11(4):262–71. doi: 10.1002/cmmi.1688
41. Feng GK, Liu RB, Zhang MQ, Ye XX, Zhong Q, Xia YF, et al. Spect and near-infrared fluorescence imaging of breast cancer with a neuropilin-1-Targeting peptide. *J Control Release* (2014) 192:236–42. doi: 10.1016/j.jconrel.2014.07.039
42. Pathak RA, Hemal AK. Intraoperative icg-fluorescence imaging for robotic-assisted urologic surgery: Current status and review of literature. *Int Urol Nephrol* (2019) 51(5):765–71. doi: 10.1007/s11255-019-02126-0
43. Hu Z, Fang C, Li B, Zhang Z, Cao C, Cai M, et al. First-in-Human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-Infrared-I/II windows. *Nat BioMed Eng* (2020) 4(3):259–71. doi: 10.1038/s41551-019-0494-0
44. Lu G, van den Berg NS, Martin BA, Nishio N, Hart ZP, van Keulen S, et al. Tumour-specific fluorescence-guided surgery for pancreatic cancer using panitumumab-IrDye800cw: A phase 1 single-centre, open-label, single-arm, dose-escalation study. *Lancet Gastroenterol Hepatol* (2020) 5(8):753–64. doi: 10.1016/s2468-1253(20)30088-1
45. Chau CH, Steeg PS, Figg WD. Antibody-drug conjugates for cancer. *Lancet* (2019) 394(10200):793–804. doi: 10.1016/s0140-6736(19)31774-x
46. Chen K, Conti PS. Target-specific delivery of peptide-based probes for pet imaging. *Adv Drug Del Rev* (2010) 62(11):1005–22. doi: 10.1016/j.addr.2010.09.004



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

YuChuan Hu,
Fourth Military Medical University, China
Wei Yang,
Beijing Cancer Hospital, Peking University,
China

*CORRESPONDENCE

Yi-Ting Yen
✉ b85401067@gmail.com
Mi-Chia Ma
✉ mcma@mail.ncku.edu.tw

[†]These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 22 November 2022

ACCEPTED 27 March 2023

PUBLISHED 18 April 2023

CITATION

Chang C-C, Tang E-K, Wei Y-F, Lin C-Y,
Wu F-Z, Wu M-T, Liu Y-S, Yen Y-T,
Ma M-C and Tseng Y-L (2023) Clinical
radiomics-based machine learning versus
three-dimension convolutional neural
network analysis for differentiation of
thymic epithelial tumors from other
prevascular mediastinal tumors on chest
computed tomography scan.
Front. Oncol. 13:1105100.
doi: 10.3389/fonc.2023.1105100

COPYRIGHT

© 2023 Chang, Tang, Wei, Lin, Wu, Wu, Liu,
Yen, Ma and Tseng. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Clinical radiomics-based machine learning versus three-dimension convolutional neural network analysis for differentiation of thymic epithelial tumors from other prevascular mediastinal tumors on chest computed tomography scan

Chao-Chun Chang^{1†}, En-Kuei Tang^{2†}, Yu-Feng Wei^{3,4†},
Chia-Ying Lin^{5†}, Fu-Zong Wu^{6,7,8}, Ming-Ting Wu^{6,9,10},
Yi-Sheng Liu⁵, Yi-Ting Yen^{1,11*}, Mi-Chia Ma^{12*}
and Yau-Lin Tseng¹

¹Division of Thoracic Surgery, Department of Surgery, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, ²Division of Thoracic Surgery, Department of Surgery, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan, ³School of Medicine for International Students, College of Medicine, I-Shou University, Kaohsiung, Taiwan, ⁴Division of Chest Medicine, Department of Internal Medicine, E-Da Cancer Hospital, Kaohsiung, Taiwan, ⁵Department of Medical Imaging, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, ⁶Department of Radiology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan, ⁷Faculty of Clinical Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, ⁸Institute of Education, National Sun Yat-sen University, Kaohsiung, Taiwan, ⁹School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, ¹⁰Institute of Clinical Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, ¹¹Division of Trauma and Acute Care Surgery, Department of Surgery, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, ¹²Department of Statistics and Institute of Data Science, National Cheng Kung University, Tainan, Taiwan

Purpose: To compare the diagnostic performance of radiomic analysis with machine learning (ML) model with a convolutional neural network (CNN) in differentiating thymic epithelial tumors (TETs) from other prevascular mediastinal tumors (PMTs).

Methods: A retrospective study was performed in patients with PMTs and undergoing surgical resection or biopsy in National Cheng Kung University Hospital, Tainan, Taiwan, E-Da Hospital, Kaohsiung, Taiwan, and Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan between January 2010 and December 2019. Clinical data including age, sex, myasthenia gravis (MG) symptoms and pathologic diagnosis were collected. The datasets were divided into UECT (unenhanced computed tomography) and CECT (enhanced computed tomography) for analysis and modelling. Radiomics model and 3D CNN model were used to differentiate TETs from non-TET PMTs (including cyst, malignant germ cell tumor, lymphoma and teratoma). The macro F1-score and receiver operating characteristic (ROC) analysis were performed to evaluate the prediction models.

Result: In the UECT dataset, there were 297 patients with TETs and 79 patients with other PMTs. The performance of radiomic analysis with machine learning model using LightGBM with Extra Tree (macro F1-Score = 83.95%, ROC-AUC = 0.9117) had better performance than the 3D CNN model (macro F1-score = 75.54%, ROC-AUC = 0.9015). In the CECT dataset, there were 296 patients with TETs and 77 patients with other PMTs. The performance of radiomic analysis with machine learning model using LightGBM with Extra Tree (macro F1-Score = 85.65%, ROC-AUC = 0.9464) had better performance than the 3D CNN model (macro F1-score = 81.01%, ROC-AUC = 0.9275).

Conclusion: Our study revealed that the individualized prediction model integrating clinical information and radiomic features using machine learning demonstrated better predictive performance in the differentiation of TETs from other PMTs at chest CT scan than 3D CNN model.

KEYWORDS

radiomics, convolutional neural networks, deep learning, machine learning, prevascular mediastinal tumor

Introduction

Prevascular mediastinal tumor (PMT) is relatively uncommon, making up less than 1% of all solid tumors (1). PMT consists of a wide variety of entities, including thymic epithelial tumor (TET), lymphoma, germ cell tumor, ectopic thyroid, and cyst, among which TET is the most frequently encountered (2). The NCCN guidelines suggested that patients with clinically resectable TETs undergo upfront surgical resection instead of preoperative transpleural biopsy to avoid converting stage I thymoma to stage IV thymoma by spreading tumor within the pleural space (3). Chest computed tomography (CT) scan is the standard assessment for PMT. It was reported that the accuracy rate of PMT interpretation *via* traditional radiographic features on CT scan reaches as high as 61% in experienced radiologists, leaving much room for improvements in the era of advanced technology (4).

The applications of radiomics in diagnostic medicine and outcome analysis have been increasingly proposed lately (5, 6). By combining image-filtering and feature-extraction methods, it is possible to extract a large number of high-order radiomic features from CT images (5). Studies have shown significant radiomic parameters such as skewness, kurtosis, and entropy, correlated with thymic tumor histology (7, 8). Redundancy is often the scenario in highly dimensional data, and classification model could be developed only through proper feature selection and proper machine learning (9).

Convolutional neural networks (CNNs) are a class of deep learning (DL) models that combine imaging filters with artificial neural networks through a series of successive linear and nonlinear layers (10, 11). CNN is far more data hungry because of its millions of learnable parameters to estimate, and therefore is more

computationally expensive, resulting in the requirement of graphical processing units (GPUs) for model training. The major drawback in the application of 3D deep learning on medical images is its dependency on data availability and high computational cost (12). With powerful GPUs becoming increasingly available, we have seen exponential growth in the applications of 3D deep learning in different medical image modalities (11).

Nonetheless, with the low incidence of tumor occurrence and the resultant limited radiographic data and information, it is yet to be clarified if 3D CNN outperforms radiomics with ML in differentiating various kinds of mediastinal tumor. Our study aimed to compare the model using radiomics with traditional machine learning with 3D CNN model in differentiating TETs from PMTs, thus providing a prediction tool and the opportunity of improvement on the decisions for invasive diagnostic or treatment modalities.

Materials and methods

Study population

A retrospective study was performed in patients with PMTs and undergoing surgical resection or biopsy in National Cheng-Kung University Hospital, Tainan, Taiwan, E-Da Hospital, Kaohsiung, Taiwan, and Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan between January 2010 and December 2019. Informed consent was waived because the study was retrospective, and it was respectively approved by the Institutional Review Board of National Cheng Kung University Hospital (A-ER-111-211), E-Da Hospital (EMRP-110-145), and Kaohsiung Veteran General

Hospital (VGHS19-CT6-08). Exclusion criteria were patients younger than 20-year-old, missing imaging data, or metastatic prevascular mediastinal tumors. The patients younger than 20-year-old were excluded because the most common types of tumors in this age group are lymphomas and malignant germ cell tumors, and the rare occurrence of thymic epithelial tumors (TET) does not have much impact on histological classification of tumors (13). Clinical data including age, sex, myasthenia gravis (MG) symptoms and pathologic diagnosis were collected. We divided our dataset into UECT (unenanced computed tomography) and CECT (enhanced computed tomography) for analysis and modelling. Radiomic model and 3D CNN model were adopted respectively to differentiate TETs from non-TET PMTs (including cyst, malignant germ cell tumor, lymphoma and teratoma). The flowchart of patient inclusion was shown in Figure 1.

Image acquisition

All CT images were obtained using Siemens SOMATOM Definition Flash, Siemens SOMATOM Definition AS, Siemens SOMATOM Sensation 16, GE Optima CT660, GE Revolution CT, GE Bright Speed Elite, GE light speed VCT, and TOSHIBA CT64-TSX-01A64. The CT protocols were as follows: 120 kVp; tube current, 150–200 mAs with automatic tube current modulation. The section thickness ranged between 0.7 mm and 1.5 mm, and the image size was 512 × 512 pixels. The detailed protocol and contrast materials are summarized in Table E1. Three patients received solely non-enhanced CT scan due to renal function impairment, while the other patients received both non-enhanced and enhanced CT scan. Contrast enhanced images were obtained after intravenous administration of contrast medium (injection dose 60–120 mL at a rate of 1.5–3 ml/s) followed by a 20 ml saline flush. Contrast enhanced images were obtained 90s after contrast agent administration. All images were reconstructed into 5-mm sections.

Tumor segmentation and image preprocessing

CT images were imported into the open-source software 3D Slicer 4.10.2 and the tumors were then contoured manually by one of two observers (C.Y.L. a radiologist with 9 years of experience and C.C.C. a thoracic surgeon with 9 years of experience) blinded to patient diagnosis using the built-in paint tool (14). The delineation of tumor at UECT and CECT was performed in the mediastinal setting (window level, 50 HU; window width, 350 HU) on the axial CT plane. Consensus was reached by discussion if the interobserver variability was apparent. For normalization, all CT voxels were resampled to 1 mm³ using a cubic interpolation.

Radiomics feature extraction, selection and model building

The global framework showing the radiomic analysis process is shown in Figure 2. The whole PMTs in each CT examination served as VOIs, from which 3D radiomic features were extracted using the open-source platform PyRadiomics (15). A total of 851 radiomic features, including 14 shape features, 18 intensity histogram features, 74 texture features, and 745 wavelet features were extracted for further analysis.

A multivariate logistic regression model was developed using the least absolute shrinkage and selection operator (LASSO) with L1 penalty to filter the features to reduce the redundancy of the features. The features with non-zero coefficients at optimized hyperparameter lambda were selected and used in ML. Relevant clinical information including age, sex and MG symptoms were also input as feature vectors in ML.

In combination with feature-selection method, eight ML classifiers were used to differentiate thymoma from other PMTs: KNeighbors, random forest (RF), extreme gradient boosting

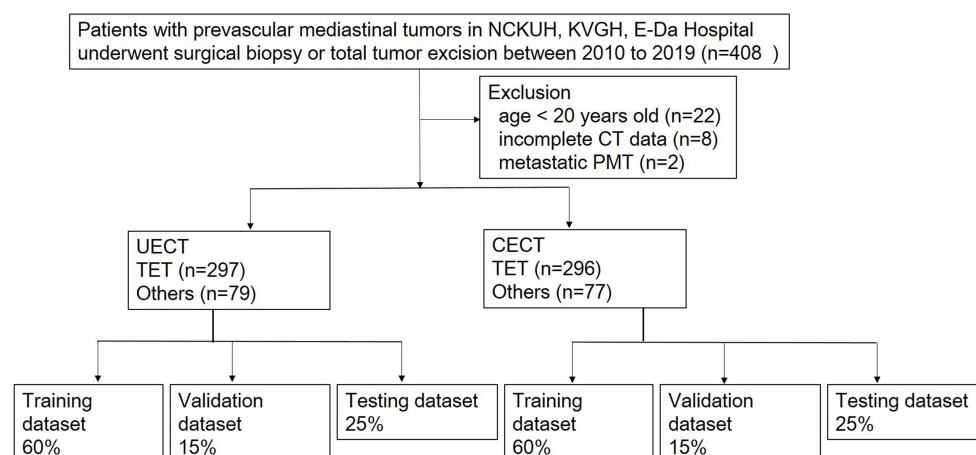


FIGURE 1

Flowchart of patients inclusion. NCKUH, National Cheng Kung University Hospital; KVGH, Kaohsiung Veteran General Hospital.

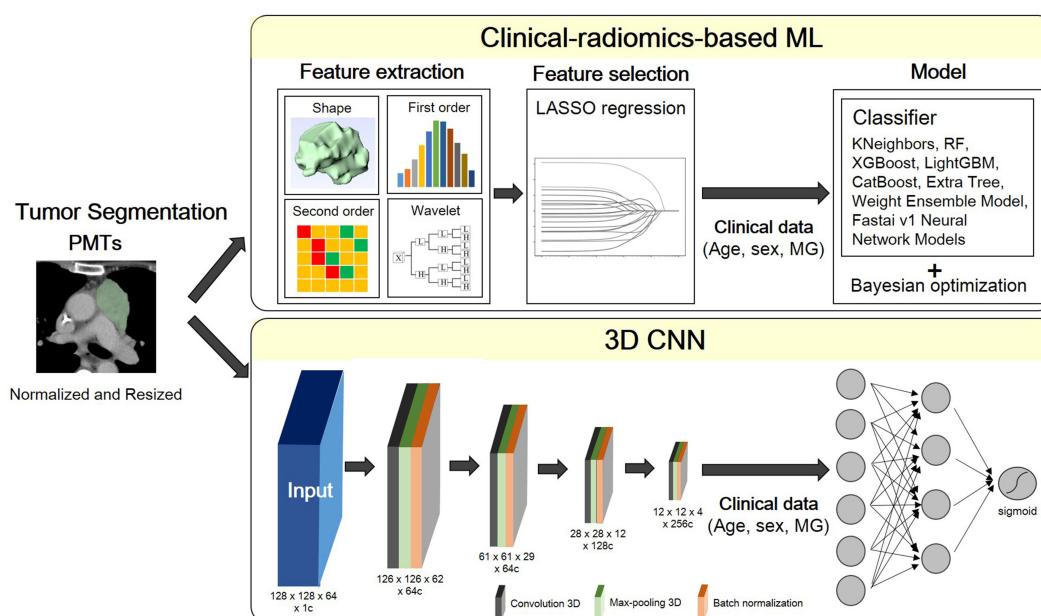


FIGURE 2

Flowchart of the proposed scheme. PMT, prevascular tumor; LASSO, least absolute shrinkage and selection operator; MG, myasthenia gravis; RF, random forest; XGBoost, extreme gradient boosting; LightGBM, light gradient boosting machine.

(XGBoost), Light Gradient Boosting Machine (LightGBM), CatBoost, Extra Tree, and Weight Ensemble Model (WeightEnsemble_L2). A Bayesian optimization algorithm (BOA) was applied to optimize the hyperparameters of these models. The flowchart of the proposed scheme was illustrated in Figure 2.

3D convolutional neural networks model

A fully convolutional neural network (CNN), as described in a recently published previous work was used (16). ROI patches were extracted from each CT image by defining a bounding box that enclosed each previously defined ROI. All patches were resized to $128 \times 128 \times 64$. Owing to the limited training data, we applied random rotation ($-30^\circ \sim 30^\circ$) for training data augmentation. As radiomics-based machine learning model, the clinical information including age, sex and MG symptoms were also input as feature vectors in this neural network model.

All the models are implemented in Python 3.8.9 based on tensorflow 2.8.0 and trained on 1 Tesla V100-DGXS-32GB. The loss function during training was the binary cross-entropy metric and was optimized using an Adam optimizer with a batch size of 2, and an initial learning rate of 10^{-5} . Early stopping was employed to prevent overfitting, and training was stopped once model performance stops improving on a validation dataset after 350 constitutive training epochs (maximum epoch value = 1000). The best model observed during training would be the output model. The flowchart of the proposed scheme was illustrated in Figure 2.

Statistical analyses

Continuous variables were compared using the Student t-test, and categorical variables were compared using the chi-square test. P values of < 0.05 were considered statistically significant.

Clinical information (age, sex and presence of MG symptom) was added into radiomic model and 3D CNN model. Each dataset was randomly split into training, validation and testing sets in the ratio of 60:15:25. The models were evaluated through repeated random sub-sampling validation.

The accuracy, macro precision, macro recall, and macro F1-score for each dataset were calculated and determined to verify the performance of the different models. The macro F1-score reflected the effectiveness on small classes and is an effective evaluation metric for an imbalanced dataset. Receiver operating characteristic (ROC) analysis was performed, and the area under the curve (AUC) was used to evaluate the prediction models. The analysis was performed using python 3.8.9 with scikit-learn 1.0.1, autogluon 0.2.0, and statsmodels 0.13.1.

Result

Basic clinical characteristics

The clinicopathological characteristics of patients in our study are shown in Table 1. In UECT dataset, there were 297 patients with TETs and 79 patients with other PMTs. In CECT dataset, there were 296 patients with TETs and 77 patients with other PMTs. There was

TABLE 1 Baseline characteristics of patients included for analysis.

Variables	UECT			CECT		
	TET (n=297)	Others (n=79)	p value	TET (n=296)	Others (n=77)	p value
Sex (male)	135 (45.5%)	36 (45.6%)	1	134 (45.3%)	36 (46.8%)	0.898
Age (y)*	61.70 ± 12.99	51.19 ± 17.74	<0.001	61.46 ± 13.06	49.26 ± 17.76	<0.001
Myasthenia gravis	87 (29.3%)	4 (5.1%)	<0.001	87 (29.4%)	4 (5.2%)	<0.001
Pathology						
Thymoma	240 (80.8%)			239 (80.7%)		
Thymic carcinoma	57 (19.2%)			57 (19.3%)		
Cyst		43 (54.4%)			41 (53.2%)	
Malignant germ cell tumor		10 (12.7%)			10 (13.0%)	
Lymphoma		11 (13.9%)			11 (14.3%)	
Teratoma		15 (19.0%)			(19.5%)	

Except where indicated, data are numbers of patients, with percentages in parentheses.

* Data are means ± SDs.

significant difference in age and MG symptoms in UECT dataset [TETs vs. other PMTs: 61.70 ± 12.99 vs. 51.19 ± 17.74, $p < 0.001$; 87 (29.3%) vs. 4 (5.1%), $p < 0.001$]. In TETs, there were 19.2% thymic carcinoma and 80.8% thymomas. In other PMTs, there were mostly cyst (54.4%). In CECT dataset, there was significant difference in age and MG symptoms [TETs vs. other PMTs: 61.46 ± 13.06 vs. 49.26 ± 17.76, $p < 0.001$; 87 (29.4%) vs. 4 (5.2%), $p < 0.001$]. In TETs, there were 19.3% thymic carcinoma and 80.7% thymomas. In other PMTs, there were mostly cyst (53.2%).

Radiomics feature selection and optimal signature construction

After performing tumor segmentation in the included patients, 851 radiomics features were extracted (Supplementary Tables 1, 2). The top five feature selection using LASSO logistic regression with different values of lambda was shown in Tables 2, 3, and the top 20 feature selection was showed in Supplementary Tables 3, 4.

The result of best feature selection using various machine learning methods was demonstrated in Table 4. In UECT dataset, LightGBM with Extra Tree using features in selection_5 had best performance (macro F1-Score = 83.95%, accuracy = 89.99%). The ROC curve was shown in Figure 3 with AUC = 0.9117. In CECT dataset, LightGBM with Extra Tree using features in selection_4 had best performance (Macro F1-Score = 85.65%, accuracy = 91.15%). The ROC curve was shown in Figure 3 with AUC = 0.9464. The results of Bayesian optimization of different models and various feature selection in UECT and CECT were showed in Supplementary Tables 5, 6.

3D CNN analysis

The result of 3D CNN classification was shown in Table 4. In UECT dataset, macro F1-score was 75.54%, accuracy 84.16%. The

ROC curve was shown in Figure 3 with AUC = 0.9015. In CECT dataset, macro F1-score was 81.01%, accuracy 86.73%. The ROC curve was shown in Figure 3 with AUC = 0.9275. Because we used repeated random sub-sampling validation method for 10 times, the total training time was 14.5 hours in UECT dataset and 14.1 hours in CECT dataset.

In comparison, the performance of radiomic analysis with machine learning model using LightGBM with Extra Tree had better performance than the 3D CNN model in both UECT and CECT dataset. Four cases were illustrated to differentiate thymic epithelial tumors from other prevascular mediastinal tumors with our ML and 3D CNN classification models in the Figure 4.

Discussion

Our result showed radiomics with ensemble machine learning achieved better performance than 3D CNN in differentiating TETs from other PMTs. Deep learning (DL) model presented more stable shape than radiomics with ML model on ROC curve. Radiomics with ML and DL are active research in the field of oncology (17). Some studies showed that the DL model had better performance than the ML-based radiomics (18–20), some showed ML-based radiomics out-performed DL model (21), and some demonstrated DL-based radiomics model had the best performance (22, 23). Prior studies had performed radiomics based ML or DL to classify thymoma from other PMT (24–26). However, our study was the first to compare the performance of radiomics-based ML with DL using the same dataset to differentiated thymoma from other PMT. While it is well-known that with large datasets, the performance of DL model was superior to hand-crafted feature extraction, a large dataset is not always available in medicine and may be limited by factors such as disease incidence, prevalence, and obstacles in data procurement. For small dataset, studies have suggested feature engineering may be a more suitable machine learning strategy with notable advantages of radiomics for medical imaging

TABLE 2 Top five variable feature selection performed by LASSO Logistic Regression at various lambda value using UECT.

Feature Selection	Lambda	Selected Variable		
Selection 1	0.04863	original_firstorder_Median	wavelet-LLH_glcmm_MCC	wavelet-LLH_glrmm_RunEntropy
		wavelet-LLH_glszm_LargeAreaLowGrayLevelEmphasis	wavelet-HLH_ngtdm_Busyness	wavelet-LLL_firstorder_RootMeanSquared
		wavelet-LLL_glszm_LargeAreaEmphasis		
Selection 2	0.03511	original_firstorder_Median	original_glrmm_RunEntropy	wavelet-LLH_glcmm_Correlation
		wavelet-LLH_glcmm_MCC	wavelet-LLH_glrmm_RunEntropy	wavelet-LLH_glrmm_ShortRunLowGrayLevelEmphasis
		wavelet-LLH_glszm_LargeAreaLowGrayLevelEmphasis	wavelet-LHL_glszm_SmallAreaEmphasis	wavelet-LHH_glrmm_RunEntropy
		wavelet-HLH_ngtdm_Busyness	wavelet-HHL_glszm_GrayLevelNonUniformityNormalized	wavelet-LLL_firstorder_RootMeanSquared
		wavelet-LLL_glszm_LargeAreaEmphasis		
Selection 3	0.02535	original_shape_Sphericity	original_firstorder_Median	original_glrmm_RunEntropy
		wavelet-LLH_glcmm_Correlation	wavelet-LLH_glcmm_MCC	wavelet-LLH_glrmm_RunEntropy
		wavelet-LLH_glrmm_ShortRunLowGrayLevelEmphasis	wavelet-LLH_glszm_LargeAreaLowGrayLevelEmphasis	wavelet-LHL_firstorder_Skewness
		wavelet-LHL_glszm_SmallAreaEmphasis	wavelet-LHH_glcmm_ClusterShade	wavelet-LHH_glrmm_RunEntropy
		wavelet-HLH_ngtdm_Busyness	wavelet-HHL_glszm_GrayLevelNonUniformityNormalized	wavelet-HHH_glszm_SizeZoneNonUniformityNormalized
		wavelet-HHH_glszm_SmallAreaEmphasis	wavelet-LLL_firstorder_RootMeanSquared	wavelet-LLL_glszm_LargeAreaEmphasis
Selection 4	0.01831	original_shape_Sphericity	original_firstorder_Median	original_glrmm_RunEntropy
		wavelet-LLH_glcmm_Correlation	wavelet-LLH_glcmm_MCC	wavelet-LLH_glrmm_RunEntropy
		wavelet-LLH_glrmm_ShortRunLowGrayLevelEmphasis	wavelet-LLH_glszm_LowGrayLevelZoneEmphasis	wavelet-LHL_glszm_SmallAreaEmphasis
		wavelet-LHH_firstorder_Mean	wavelet-LHH_glcmm_MaximumProbability	wavelet-LHH_glrmm_RunEntropy
		wavelet-LHH_ngtdm_Busyness	wavelet-HLL_glszm_ZoneVariance	wavelet-HLH_glcmm_MCC
		wavelet-HLH_glszm_SmallAreaLowGrayLevelEmphasis	wavelet-HLH_ngtdm_Busyness	wavelet-HHL_glszm_GrayLevelNonUniformityNormalized
		wavelet-HHH_glszm_SizeZoneNonUniformityNormalized	wavelet-HHH_glszm_SmallAreaEmphasis	wavelet-LLL_firstorder_RootMeanSquared
Selection 5	0.01322	original_shape_Sphericity	original_shape_SurfaceVolumeRatio	original_firstorder_Median
		original_firstorder_c	original_glrmm_RunEntropy	wavelet-LLH_firstorder_Kurtosis
		wavelet-LLH_glcmm_Correlation	wavelet-LLH_glcmm_MCC	wavelet-LLH_glcmm_MaximumProbability
		wavelet-LLH_glrmm_ShortRunLowGrayLevelEmphasis	wavelet-LLH_glszm_LowGrayLevelZoneEmphasis	wavelet-LHL_firstorder_Skewness
		wavelet-LHL_glszm_SmallAreaEmphasis	wavelet-LHH_firstorder_Mean	wavelet-LHH_glcmm_MaximumProbability
		wavelet-LHH_glrmm_RunEntropy	wavelet-LHH_glszm_SizeZoneNonUniformityNormalized	wavelet-LHH_ngtdm_Busyness

(Continued)

TABLE 2 Continued

Feature Selection	Lambda	Selected Variable		
		wavelet-HLL_firstorder_Maximum	wavelet-HLL_glszm_ZoneVariance	wavelet-HLH_glcM_MCC
		wavelet-HLH_glcM_MaximumProbability	wavelet-HLH_glszm_SmallAreaLowGrayLevelEmphasis	wavelet-HLH_ngtdm_Busyness
		wavelet-HHL_glcM_Idmn	wavelet-HHL_glszm_GrayLevelNonUniformityNormalized	wavelet-HHH_glrIm_ShortRunLowGrayLevelEmphasis
		wavelet-HHH_glszm_SizeZoneNonUniformityNormalized	wavelet-HHH_glszm_SmallAreaEmphasis	wavelet-LLL_firstorder_RootMeanSquared
		wavelet-LLL_glcM_MaximumProbability		

TABLE 3 Top five variable feature selection performed by Lasso Logistic Regression at various lambda value using CECT.

Feature Selection	Lambda	Selected Variable		
Selection 1	0.04863	wavelet-LHH_firstorder_Mean	wavelet-LHH_glcM_SumEntropy	wavelet-LHH_gldm_LargeDependenceEmphasis
		wavelet-LHH_ngtdm_Busyness	wavelet-HLH_glszm_SizeZoneNonUniformityNormalized	wavelet-HLH_ngtdm_Busyness
		wavelet-LLL_firstorder_Median		
Selection 2	0.03511	wavelet-LHL_firstorder_Mean	wavelet-LHH_firstorder_Mean	wavelet-LHH_glcM_SumEntropy
		wavelet-LHH_glrIm_RunLengthNonUniformityNormalized	wavelet-LHH_glrIm_RunPercentage	wavelet-LHH_ngtdm_Busyness
		wavelet-HLL_glszm_LargeAreaHighGrayLevelEmphasis	wavelet-HLH_glszm_SizeZoneNonUniformityNormalized	wavelet-HLH_ngtdm_Busyness
		wavelet-HHL_glcM_Imc1	wavelet-HHL_glcM_InverseVariance	wavelet-LLL_firstorder_Median
Selection 3	0.02535	wavelet-LHL_firstorder_Mean	wavelet-LHL_firstorder_Skewness	wavelet-LHH_firstorder_Mean
		wavelet-LHH_glcM_JointEnergy	wavelet-LHH_glcM_SumEntropy	wavelet-LHH_glrIm_RunPercentage
		wavelet-LHH_glszm_SizeZoneNonUniformityNormalized	wavelet-LHH_ngtdm_Busyness	wavelet-HLL_firstorder_90Percentile
		wavelet-HLL_glszm_LargeAreaHighGrayLevelEmphasis	wavelet-HLH_glszm_SizeZoneNonUniformityNormalized	wavelet-HLH_ngtdm_Busyness
		wavelet-HHL_glcM_Imc1	wavelet-HHL_glcM_InverseVariance	wavelet-HHL_glcM_MCC
		wavelet-LLL_firstorder_Median		
Selection 4	0.01831	wavelet-LLH_glszm_GrayLevelVariance	wavelet-LLH_glszm_LowGrayLevelZoneEmphasis	wavelet-LHL_firstorder_Mean
		wavelet-LHL_firstorder_Skewness	wavelet-LHL_glcM_InverseVariance	wavelet-LHH_firstorder_Mean
		wavelet-LHH_glcM_Imc2	wavelet-LHH_glcM_JointEnergy	wavelet-LHH_glcM_SumEntropy
		wavelet-LHH_glrIm_RunLengthNonUniformityNormalized	wavelet-LHH_glrIm_RunPercentage	wavelet-LHH_glszm_SizeZoneNonUniformityNormalized

(Continued)

TABLE 3 Continued

Feature Selection	Lambda	Selected Variable		
		wavelet-LHH_ngtdm_Busyness	wavelet-HLL_firstorder_90Percentile	wavelet-HLL_glcm_Correlation
		wavelet-HLH_glszm_SizeZoneNonUniformityNormalized	wavelet-HLH_ngtdm_Busyness	wavelet-HHL_glcm_Imc1
		wavelet-HHL_glcm_InverseVariance	wavelet-HHL_glcm_MCC	wavelet-HHL_glszm_LargeAreaHighGrayLevelEmphasis
		wavelet-HHL_glszm_SmallAreaLowGrayLevelEmphasis	wavelet-LLL_firstorder_Median	
Selection 5	0.01322	original_shape_Sphericity	wavelet-LLH_glszm_GrayLevelVariance	wavelet-LLH_glszm_LowGrayLevelZoneEmphasis
		wavelet-LHL_firstorder_Maximum	wavelet-LHL_firstorder_Mean	wavelet-LHL_firstorder_Median
		wavelet-LHL_firstorder_Skewness	wavelet-LHL_glcm_InverseVariance	wavelet-LHH_firstorder_Mean
		wavelet-LHH_glcm_Imc2	wavelet-LHH_glcm_JointEnergy	wavelet-LHH_glcm_SumEntropy
		wavelet-LHH_glrml_RunLengthNonUniformityNormalized	wavelet-LHH_glrml_RunPercentage	wavelet-LHH_glszm_SizeZoneNonUniformityNormalized
		wavelet-LHH_ngtdm_Busyness	wavelet-HLL_firstorder_90Percentile	wavelet-HLL_glcm_Correlation
		wavelet-HLH_glcm_MCC	wavelet-HLH_glrml_LongRunLowGrayLevelEmphasis	wavelet-HLH_glszm_GrayLevelNonUniformityNormalized
		wavelet-HLH_glszm_SizeZoneNonUniformityNormalized	wavelet-HLH_ngtdm_Busyness	wavelet-HHL_glcm_Imc1
		wavelet-HHL_glcm_InverseVariance	wavelet-HHL_glcm_MCC	wavelet-HHL_glcm_MaximumProbability
		wavelet-HHL_glszm_LargeAreaHighGrayLevelEmphasis	wavelet-HHL_glszm_SmallAreaLowGrayLevelEmphasis	wavelet-HHH_firstorder_Median
		wavelet-LLL_firstorder_Median		

TABLE 4 The best result of radiomics models and 3D CNN model to differentiate thymoma from other prevascular mediastinal tumors.

	Feature selection	Macro F1-Score	Macro Precision	Macro Recall	Accuracy
UECT					
CatBoost	Selection 4	0.8121	0.8686	0.7813	0.8882
ExtraTrees with Entropy	Selection 1	0.7921	0.8519	0.7606	0.8786
ExtraTrees with Gini	Selection 1	0.7985	0.8595	0.7663	0.8818
Kneighbors with Distance weights	All feature	0.6128	0.6599	0.6015	0.7861
Kneighbors with Uniform weights	All feature	0.6135	0.6623	0.6015	0.7861
LightGBM	Selection 1	0.8258	0.8586	0.8052	0.8914
LightGBM with Extra Trees	Selection 5	0.8395	0.8773	0.816	0.8999
NeuralNetFastAI	Selection 9	0.833	0.835	0.8376	0.8851
Random Forest with Entropy	Selection 8	0.7952	0.8569	0.7631	0.8797
Random Forest with Gini	Selection 4	0.7984	0.8567	0.7685	0.8797
WeightedEnsemble_L2	Selection 9	0.8328	0.8605	0.8145	0.8946
XGBoost	Selection 1	0.8061	0.8305	0.7891	0.8775
3D CNN		0.7554	0.7679	0.7531	0.8416
CECT					
CatBoost	Selection 4	0.8374	0.8824	0.8111	0.9027
ExtraTrees with Entropy	Selection 1	0.796	0.8747	0.7595	0.8856
ExtraTrees with Gini	Selection 1	0.7981	0.8781	0.7602	0.8867
Kneighbors with Distance weights	All feature	0.6344	0.7036	0.6189	0.8059
Kneighbors with Uniform weights	All feature	0.6257	0.6865	0.6129	0.7985
LightGBM	Selection 2	0.8533	0.8806	0.836	0.9081
LightGBM with Extra Trees	Selection 4	0.8565	0.8889	0.8353	0.9115
NeuralNetFastAI	Selection 7	0.8388	0.8351	0.8564	0.8854
Random Forest with Entropy	Selection 3	0.8328	0.8932	0.8017	0.9039
Random Forest with Gini	Selection 3	0.8344	0.8892	0.8029	0.9027
WeightedEnsemble_L2	Selection 5	0.8506	0.8763	0.8341	0.9072
XGBoost	Selection 3	0.8496	0.8601	0.8432	0.9029
3D CNN		0.8101	0.8178	0.8082	0.8673

analysis. At present, studies that directly compare radiomics and deep learning clinical model performance are relatively unexplored. In this study, we address these questions and further aim to enhance interpretability of such machine learning models (27).

There are a large number of vectors and associated computational cost in DL. We used 3D CNN in our study, which was relatively simple; however, the trainable parameters were up to 1,351,873. Because of the long training time, it would be of more difficulty for researchers to modify the algorithm repeatedly. In comparison, radiomics with ML had lower computational cost but was associated with more complicated process. After radiomic feature extraction, most feature vectors are redundant. Therefore, feature selection is demanded to build a model *via* ML. The method

of ML has substantial impact on its performance. In our study, not all radiomics with ML method out-performed 3D CNN. In consistent with prior study, ensemble learning had the best result. Ensemble learning can also be applied in DL, with the cost of longer training time.

Our results revealed that dataset of CECT worked better than UECT in classifying thymoma from other PMTs using both radiomics models and 3D CNN model, which was consistent with our clinical experience. The imaging characters at chest CT scan of low-risk thymoma and thymic cyst showed round or oval shape, smooth contour, while high-risk thymoma showed irregular shape and contour. Nonseminomatous germ cell tumor demonstrated marked hemorrhagic necrosis, while teratoma revealed fat

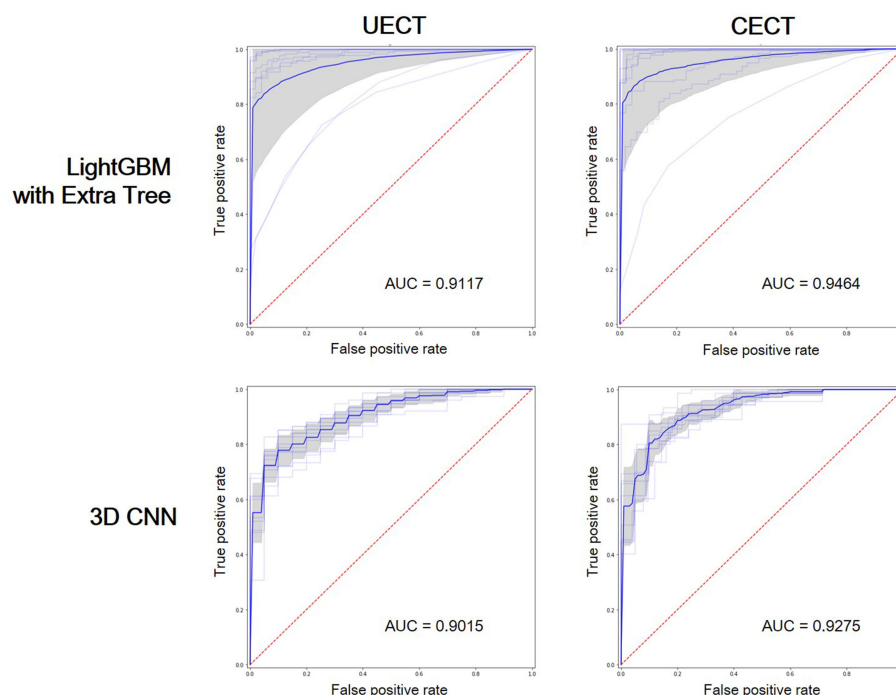


FIGURE 3

Receiver operating characteristic (ROC) curves showed the performance of LightGBM with Extra Tree and 3D CNN on UECT and CECT. AUC, area under curve.

component (2). From the result of LASSO selection in UECT and CECT, sphericity in shape feature played an important role in two dataset model, consistent with the finding in our conventional CT scan. High resolution medical imaging contains many features that is difficult to discover by visual inspection. The ability of multi-scale and multiresolution in wavelet transform has been verified in many imaging studies, and often applied to image compression, edge detection, feature extraction, and texture analysis. Our study demonstrated that wavelet-based features were selected by two datasets, suggesting the importance of high order features in imaging identification. However, other shape features and original first order features are more important in classification in the UECT dataset than in the CECT dataset. Compared with CECT, UECT lacks contrast agent to demonstrate richer texture features of soft tissue, and the septa within the tumor or the range of necrosis are less clearly seen. This could be responsible for the reason that UECT had a different tendency of feature selection from CECT dataset.

In differentiating TET from cysts, radiologists primarily focus on the Hounsfield units (HU) changes between non-enhanced and contrast-enhanced scans. Previous studies have found that cysts have a mean attenuation value of around 23 HU and a maximal attenuation value of 58 HU (28). However, some thymic cysts may have increased CT attenuation if hemorrhage or infection occurs, and relying solely on non-enhanced scans may lead to misdiagnosis. According to our research approach, we have developed separate models for predicting TET from non-enhanced and contrast-enhanced CT scans. Interestingly, although contrast-enhanced CT has better predictive performance, using non-enhanced CT alone, whether based on radiomic-based machine learning or 3D

convolutional neural network, achieves an AUC of > 0.9, which is close to the performance of the contrast-enhanced CT group. The difference of macro F1-score in all ML methods and 3D CNN in both UECT and CECT datasets was less than five percent. Tumor segmentation in UECT is sometimes difficult due to its proximity to adjacent vessel, heart, pericardial effusion, or consolidated lung. However, radiomics-based ML achieved an accuracy of 90%, indicating that our model had good performance using UECT. As the LDCT for lung cancer screening become more prevalent, there is increasing number of incidentally found asymptomatic PMTs. Once the UECT dataset provides high accuracy in differentiating PMTs, patients do not need to undergo CECT, and radiation dose and contrast agent exposure with the likelihood of kidney injury could therefore be minimized. Besides, previous literature has reported that approximately 22% to 68% of non-therapeutic thymectomies were unnecessary (28). In our dataset, 14% (53/376) of the patients had cysts or lymphoma and could have been otherwise managed instead of being operated on. Our model achieved an accuracy of 0.91. Therefore, we believe that increasing the accuracy of preoperative imaging diagnosis will help to reduce unnecessary invasive procedures.

Our study had several limitations. First, tumors were manually segmented; this could be user-dependent, time-consuming and labor-intensive. Prior studies have demonstrated DL-based tumor segmentation algorithm with robust performance. Automated tumor segmentation could probably be integrated into an automated processing pipeline to minimize subjectivity and facilitate large-scale studies. Second, after radiomics feature extraction, we used only LASSO regression with variable lambda value for feature selection.

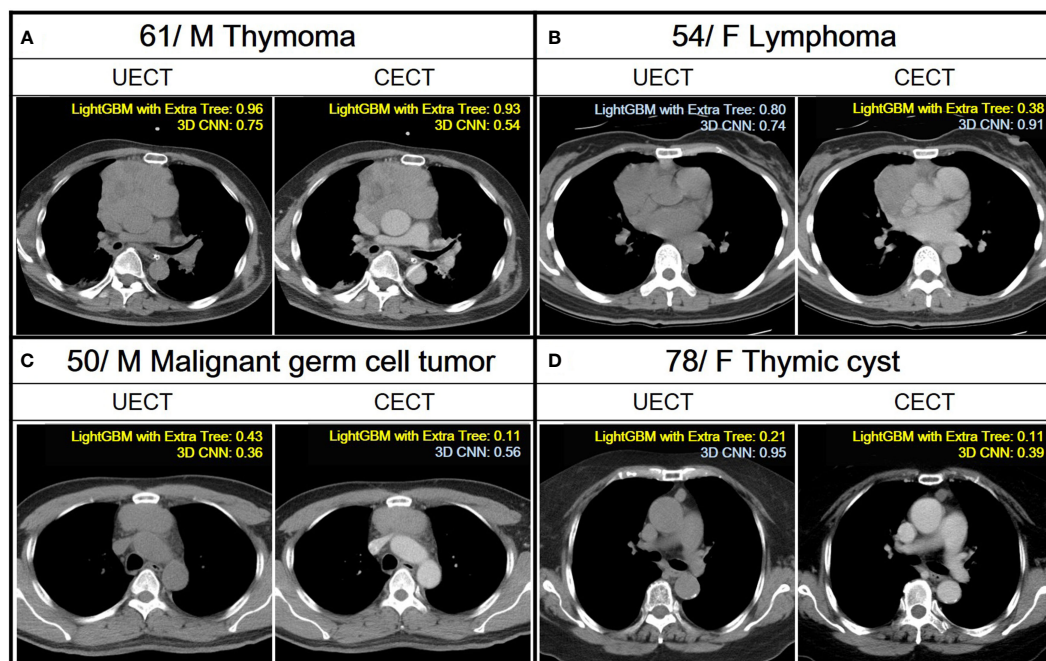


FIGURE 4

Demonstration of the application of ML and 3D CNN classification models to analyze four distinct cases with varying pathologies. A case is classified as TET if the confidence score obtained from a ML model and 3D CNN using their UECT and CECT, respectively, is greater than 0.50, and classified as non-TET if it is less than or equal to 0.50. Yellow text indicates correct predictions, while blue text indicates incorrect predictions. (A) A 61 years old male with thymoma. Both LightGBM with Extra Tree and 3D CNN had correct prediction from UECT and CECT. (B) A 54 years old female with lymphoma. The LightGBM with ExtraTree had correct prediction from CECT. (C) A 50 years old male with malignant germ cell tumor. The LightGBM with Extra Tree had correct prediction from both UECT and CECT, while 3D CNN had correct prediction from UECT. (D) A 78 years old female with thymic cyst. The LightGBM with Extra Tree had correct prediction from both UECT and CECT, while 3D CNN had correct prediction from CECT.

Adopting different feature selection methods could result in different outcome. Third, we used a relatively simple 3D CNN model for classification. Although complicated model is computationally expensive, future advancement in hardware (e.g., GPU or cloud computing) and algorithmic development is expectable. Fourth, because of the retrospective nature of this analysis, a selection bias was unavoidable. Lastly, CT images in our study were obtained using heterogeneous CT scanners, with various acquisition parameters, which can affect radiomic features and analysis. Nonetheless, the diagnostic performance of the radiomics model remained high in the validation cohort, which verified the good generalizability of the model.

As DL gradually becomes the mainstream of imaging study, some studies showed DL was superior to radiomics with ML in visual classification. However, it is susceptible to overfitting and takes a large number of data for model training and parameter tuning. Owing to the limited size of dataset, our proposed radiomics-based ML and 3D CNN scheme may be overfitting during training process. In rare disease with limited case number, radiomics-based ML may have better efficacy with lower computational cost. The best method of computing depends on the subject of study and size of dataset. Further studies are mandatory to evaluate the efficacy of ML and DL in the same dataset.

Conclusion

To conclude, an ensemble ML method with radiomic feature can be useful for differentiating TETs from other type of PMTs, performed slightly better than a 3D CNN, and demonstrated good generalizability across institutions.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by National Cheng Kung University Hospital (A-ER-111-211), E-Da Hospital (EMRP-110-145), and Kaohsiung Veteran General Hospital (VGHKS19-CT6-08). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

(I) Conception and design: C-CC, Y-TY. (II) Administrative support: None. (III) Provision of study materials or patients: C-CC, E-KT, Y-FW, C-YL, F-ZW, M-TW, Y-SL, Y-TY and Y-LT. (IV) Collection and assembly of data: C-CC, E-KT, Y-FW, F-ZW and C-YL. (V) Data analysis and interpretation: C-CC, C-YL and M-CM. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Cheng Kung University Hospital of Taiwan [NCKUH-11103026 and NCKUH-11201007], the Ministry of Science and Technology of Taiwan [MOST 110-2314-B-006-103 and MOST 111-2314-B-006-106], and Higher Education Sprout Project, Ministry of Education to the Headquarters of University Advancement at National Cheng Kung University(NCKU).

References

- Araki T, Nishino M, Gao W, Dupuis J, Washko GR, Hunninghake GM, et al. Anterior mediastinal masses in the framingham heart study: Prevalence and CT image characteristics. *Eur J Radiol Open* (2015) 2:26–31. doi: 10.1016/j.ejro.2014.12.003
- Nakazono T, Yamaguchi K, Egashira R, Mizuguchi M, Irie H. Anterior mediastinal lesions: CT and MRI features and differential diagnosis. *Japanese J Radiol* (2021) 39(2):101–17. doi: 10.1007/s11604-020-01031-2
- Network, National Comprehensive Cancer. *Thymomas and thymic carcinomas* (2022). Available at: https://www.nccn.org/professionals/physician_gls/pdf/thymic.pdf.
- Tomiya N, Honda O, Tsubamoto M, Inoue A, Sumikawa H, Kuriyama K, et al. Anterior mediastinal tumors: diagnostic accuracy of CT and MRI. *Eur J Radiol* (2009) 69(2):280–8. doi: 10.1016/j.ejrad.2007.10.002
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
- Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. *Theranostics* (2019) 9(5):1303–22. doi: 10.7150/thno.30309
- Yasaka K, Akai H, Nojima M, Shinzaki-Ushiku A, Fukayama M, Nakajima J, et al. Quantitative computed tomography texture analysis for estimating histological subtypes of thymic epithelial tumors. *Eur J Radiol* (2017) 92:84–92. doi: 10.1016/j.ejrad.2017.04.017
- Iannarelli A, Sacconi B, Tomei F, Anile M, Longo F, Bezzi M, et al. Analysis of CT features and quantitative texture analysis in patients with thymic tumors: correlation with grading and staging. *Radiol Med* (2018) 123(5):345–50. doi: 10.1007/s11547-017-0845-4
- Rizzo S, Botta F, Raimondi S, Origi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp* (2018) 2(1):36. doi: 10.1186/s41747-018-0068-z
- Lee JG, Jun S, Cho Y, W, Lee H, Kim GB, Seo JB, et al. Deep learning in medical imaging: General overview. *Korean J Radiol* (2017) 18(4):570–84. doi: 10.3348/kjr.2017.18.4.570
- Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyas B. 3D deep learning on medical images: A review. *Sensors (Basel)* (2020) 20(18):5097. doi: 10.3390/s20185097
- Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* (2018) 9(4):611–29. doi: 10.1007/s13244-018-0639-9
- Biko DM, Lichtenberger JP. 3rd, Rapp JB, Khwaja A, Huppmann AR, Chung EM. Mediastinal masses in children: Radiologic-pathologic correlation. *RadioGraphics* (2021) 41(4):1186–207. doi: 10.1148/rg.2021200180
- Court LE, Fave X, Mackin D, Lee J, Yang J, Zhang L. Computational resources for radiomics. *Trans Cancer Res* (2016) 5(4):340–8. doi: 10.21037/tcr.2016.06.17
- van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.CAN-17-0339
- Zunair H, Rahman A, Mohammed N, Cohen JP. Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction. (2020). doi: 10.1007/978-3-030-59354-4_15
- Choi JY. Radiomics and deep learning in clinical imaging: What should we do? *Nucl Med Mol Imaging* (2018) 52(2):89–90. doi: 10.1007/s13139-018-0514-0
- Abrol A, Fu Z, Salman M, Silva R, Du Y, Plis S, et al. Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nat Commun* (2021) 12(1):353. doi: 10.1038/s41467-020-20655-6
- Park YR, Kim YJ, Ju W, Nam K, Kim S, Kim KG. Comparison of machine and deep learning for the classification of cervical cancer based on cervicography images. *Sci Rep* (2021) 11(1):16143. doi: 10.1038/s41598-021-95748-3
- Xu Q, Zhu Q, Liu H, Chang L, Duan S, Dou W, et al. Differentiating benign from malignant renal tumors using T2- and diffusion-weighted images: A comparison of deep learning and radiomics models versus assessment from radiologists. *J Magn Reson Imaging* (2022) 55(4):1251–9. doi: 10.1002/jmri.27900
- Castillo TJ, Arif M, Starmans MPA, Niessen WJ, Bangma CH, Schoots IG, et al. Classification of clinically significant prostate cancer on multi-parametric MRI: A validation study comparing deep learning and radiomics. *Cancers (Basel)* (2021) 14(1):12. doi: 10.3390/cancers14010012
- Xia X, Gong J, Hao W, Yang T, Lin Y, Wang S, et al. Comparison and fusion of deep learning and radiomics features of ground-glass nodules to predict the invasiveness risk of stage-I lung adenocarcinomas in CT scan. *Front Oncol* (2020) 10:418. doi: 10.3389/fonc.2020.00418
- Astaraki M, Yang G, Zakko Y, Toma-Dasu I, Smedby O, Wang C. A comparative study of radiomics and deep-learning based methods for pulmonary nodule malignancy prediction in low dose CT images. *Front Oncol* (2021) 11:737368. doi: 10.3389/fonc.2021.737368
- Wang X, You X, Zhang L, Huang D, Aramini B, Shabatur L, et al. A radiomics model combined with XGBoost may improve the accuracy of distinguishing between mediastinal cysts and tumors: a multicenter validation analysis. *Ann Transl Med* (2021) 9(23):1737. doi: 10.21037/atm-21-5999
- Kirienco M, Ninatti G, Cozzi L, Voulaz E, Gennaro N, Barajon I, et al. Computed tomography (CT)-derived radiomic features differentiate prevascular mediastinum masses as thymic neoplasms versus lymphomas. *Radiol Med* (2020) 125(10):951–60. doi: 10.1007/s11547-020-01188-w
- Sui H, Liu L, Li X, Zuo P, Cui J, Mo Z. CT-based radiomics features analysis for predicting the risk of anterior mediastinal lesions. *J Thorac Dis* (2019) 11(5):1809–18. doi: 10.21037/jtd.2019.05.32
- Wang H, Wang L, Lee EH, Zheng J, Zhang W, Halabi S, et al. Decoding COVID-19 pneumonia: comparison of deep learning and radiomics CT image signatures. *Eur J Nucl Med Mol Imaging* (2021) 48(5):1478–86. doi: 10.1007/s00259-020-05075-4
- Ackman JB, Verzosa S, Kovach AE, Louissaint A, Jr, Lanuti M, Wright CD, et al. High rate of unnecessary thymectomy and its cause. can computed tomography distinguish thymoma, lymphoma, thymic hyperplasia, and thymic cysts? *Eur J Radiol* (2015) 84(3):524–33. doi: 10.1016/j.ejrad.2014.11.042

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1105100/full#supplementary-material>



OPEN ACCESS

EDITED BY

Aaron Fenster,
Western University, Canada

REVIEWED BY

Jaron Chong,
London Health Sciences Centre, Canada
Ana Valea,
Iuliu Hatieganu University of Medicine and
Pharmacy, Romania

*CORRESPONDENCE

Lijing Wang

✉ wanglj844@zjcc.org.cn

Lei Xu

✉ xulei@qsmm.org.cn

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 03 January 2023

ACCEPTED 14 April 2023

PUBLISHED 28 April 2023

CITATION

Wang L, Wang Y, Lu W, Xu D, Yao J,
Wang L and Xu L (2023) Differential
regional importance mapping for thyroid
nodule malignancy prediction with
potential to improve needle aspiration
biopsy sampling reliability.
Front. Oncol. 13:1136922.
doi: 10.3389/fonc.2023.1136922

COPYRIGHT

© 2023 Wang, Wang, Lu, Xu, Yao, Wang and
Xu. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Differential regional importance mapping for thyroid nodule malignancy prediction with potential to improve needle aspiration biopsy sampling reliability

Liping Wang^{1†}, Yuan Wang^{2†}, Wenliang Lu², Dong Xu^{1,3},
Jincao Yao¹, Lijing Wang^{1*} and Lei Xu^{3,4*}

¹Department of Ultrasonography, The Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer, Chinese Academy of Sciences, Hangzhou, China, ²School of Mathematical Sciences, Zhejiang University, Hangzhou, China, ³Department of Ultrasound, Zhejiang Society for Mathematical Medicine, Hangzhou, China, ⁴Group of Computational Imaging and Digital Medicine, Zhejiang Qiushi Institute for Mathematical Medicine, Hangzhou, China

Objective: Existing guidelines for ultrasound-guided fine-needle aspiration biopsy lack specifications on sampling sites, but the number of biopsies improves diagnostic reliability. We propose the use of class activation maps (CAMs) and our modified malignancy-specific heat maps that locate important deep representations of thyroid nodules for class predictions.

Methods: We applied adversarial noise perturbations to the segmented concentric “hot” nodular regions of equal sizes to differentiate regional importance for the malignancy diagnostic performances of an accurate ultrasound-based artificial intelligence computer-aided diagnosis (AI-CADx) system using 2,602 retrospectively collected thyroid nodules with known histopathological diagnosis.

Results: The AI system demonstrated high diagnostic performance with an area under the curve (AUC) value of 0.9302 and good nodule identification capability with a median dice coefficient >0.9 when compared to radiologists’ segmentations. Experiments confirmed that the CAM-based heat maps reflect the differentiable importance of different nodular regions for an AI-CADx system to make its predictions. No less importantly, the hot regions in malignancy heat maps of ultrasound images in comparison with the inactivated regions of the same 100 malignant nodules randomly selected from the dataset had higher summed frequency-weighted feature scores of 6.04 versus 4.96 rated by radiologists with more than 15 years of ultrasound examination experience according to widely used ultrasound-based risk stratification American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) in terms of nodule composition, echogenicity, and echogenic foci, excluding shape and margin attributes, which could only be evaluated on the whole rather than on the sub-nodular component levels. In addition, we show examples

demonstrating good spatial correspondence of highlighted regions of malignancy heat map to malignant tumor cell-rich regions in hematoxylin and eosin-stained histopathological images.

Conclusion: Our proposed CAM-based ultrasonographic malignancy heat map provides quantitative visualization of malignancy heterogeneity within a tumor, and it is of clinical interest to investigate in the future its usefulness to improve fine-needle aspiration biopsy (FNAB) sampling reliability by targeting potentially more suspicious sub-nodular regions.

KEYWORDS

tumor heterogeneity, ultrasound guided biopsy, class activation map, artificial intelligence, thyroid nodule

Introduction

Thyroid nodules are detected in as high as 65% of the population (1), and thyroid cancer is one of the most frequently occurring malignant tumors. Meanwhile, ultrasound imaging is the most commonly used method for evaluating thyroid nodules, given its ease to detect the nodules and good sensitivity to differentiate benign from malignant tumors as well as its non-invasive nature with widespread accessibility in clinics. However, the diagnosis of thyroid nodules is highly dependent on radiologists' personal experience and subjective judgment, leading to not uncommonly inconsistent conclusions. Currently, cytopathological examination performed on minimally invasive fine-needle aspiration biopsy (FNAB) typically has a diagnostic sensitivity and specificity that vary at 65%–98% and 72%–100%, respectively (2–4). More importantly, as cancerous masses are typically heterogeneous (5, 6), it is of clinical importance to be able to differentiate regions of different malignancy levels within the same imaged tissue such that FNAB sampling can be more precisely guided. To date, there are numerous practical guidelines about under what circumstances FNAB shall be applied (7–9), but there exists no consensus or guideline about the number of needle-sampling passes to acquire adequate specimens for diagnostic purposes, let alone recommendations for precise sampling sites in nodules under ultrasound guidance. With the advancement of artificially intelligent technologies and especially the development of deep learning algorithms, it becomes increasingly common for radiologists to include these auxiliary mathematical models in their toolboxes during clinical studies for disease detection and diagnosis (10–12).

Surprisingly or not, the capability of convolutional neural networks (CNNs) to approximate any arbitrary functions has become a double-edged sword, as any insights about how the models come to their conclusions are hardly accessible to human understanding given their architectural designs. Some proposed circumventing strategies include weighting the predicted classification probability of a CNN model as a contributing factor together with other human-interpretable image features (13) and

taking the similarities between imaged lesions with known diagnosis as an extra channel of information to guide human-centered diagnosis (14).

Other possibilities to retrieve hints using computer-aided artificial intelligence diagnosis as we perceive include generating the class activation map (CAM) (15) that localizes the deep representation of class-discriminating image regions. In the field of medical imaging, CAMs have been employed to visualize hot regions that conclude each predicted classification type of tissue (16–18) using a heat map representation. To date, it is however mostly a visualization tool and lacks quantitative validation to show whether hot regions in such heat maps for malignant samples indeed possess higher importance in determining the classification type. If so, the question of how much more important the hot regions are relative to other regions still awaits an answer.

To answer these questions, we designed this proof-of-concept computational study described as the following. We first generated CAMs in two different rendering configurations using the Software Development Kit (SDK) of the artificial intelligence computer-aided diagnosis (AI-CADx) system referred to as “AI-SONIC™ Thyroid” for thyroid nodule diagnoses. In the first configuration, the heat maps were rendered conventionally such that they make no distinctions between benign and malignant cases visually, and the color temperature is supposed to show the associated regional importance in predicting the classification regardless of what the predicted type is. In the second configuration, however, the intensities of these heat maps were normalized to the malignancy probabilities predicted by the AI-CADx system and rendered in such a way that the more reddish the color, the higher probability that the nodule is predicted to be malignant, whereas the bluish color indicates benignity. In other words, the second visualization configuration presents essentially malignancy heat maps.

To quantitatively analyze the importance of different regions within the ultrasound-imaged thyroid nodules in the diagnosis by the AI-CADx system, we subdivided each nodular CAM into five concentric areas of the same size and then evaluated the AI diagnostic performances after the adversarial noises (19) were applied to each subdivided nodular region in ultrasound images.

The motivation behind this is that, hypothetically, the regions of higher importance for predicting the correct diagnosis of thyroid nodules should be more vulnerable to noise perturbations. As a second attempt, we also performed a test by varying the heat intensity threshold to segment CAMs and evaluated how the diagnoses were based on regions above the thresholds.

Materials and methods

Data acquisition

The ultrasound images covering a total of 2,602 thyroid nodules from 2,488 patients were collected between January 2011 and February 2019 by The Cancer Hospital of the University of Chinese Academy of Sciences in Hangzhou, Zhejiang, China. All nodules were diagnosed by surgical pathological examinations, among which 1,581 cases were determined to be benign and 1,021 cases were found to be malignant. The local ethics committee waived the ethical approval in view of the retrospective nature of the study, and all reviews of the ultrasound image and postsurgical hematoxylin and eosin (H&E)-stained pathological images being performed were part of the clinical routine.

AI-CADx model

This study was based on the AI-SONICTM Thyroid nodule diagnosis system (Demetics Medical Technology, Hangzhou, China). It is built on the EfficientNet (20) architectural backbone and employs supervised sharpness-aware minimization for model parameter optimization to realize automatic nodule segmentation and classification (21). The classification module of the AI model relies on the precise localization of the nodules. This system can automatically detect thyroid nodules in two-dimensional grayscale ultrasound images and output corresponding masks. Therefore, radiologists do not need to manually outline the thyroid nodules except on rare occasions when manual corrections are necessary.

Heat map

All the generated heat maps in this study were based on CAM (15) using global average pooling (GAP) in CNNs. Before the final output layer for image classification, we performed global average pooling on the convolutional feature maps and used those as features for a fully connected layer that produces the desired output. We projected back the weights of the output layer onto the convolutional feature maps using the cited CAM method. Global average pooling outputs the spatial average of the feature map of each unit at the last convolutional layer. A weighted sum of these values is used to generate the final output. We computed a weighted sum of the feature maps of the last convolutional layer to obtain our class activation maps.

For an image, we used $g_k(x, y)$ to represent the activation of unit k in the last convolutional layer at the spatial location. Next, for unit (x, y) , G^k is the result of performing global average pooling, which is defined as $\sum_{x,y} g_k(x, y)$. Then, for a class i , the input to the softmax S_i is $\sum_k \omega_k^i G_k$, where the ω_k^i is weight corresponding to class i for unit k . In particular, the ω_k^i implies the importance of class i . The final output of the softmax for class i , P_i is given by $\frac{\exp(S_i)}{\sum_i \exp(S_i)}$. The bias term was ignored as in the original paper by setting the input bias of the softmax to 0 to have no impact on the classification performance.

By adding $G_k = \sum_{x,y} g_k(x, y)$ into the class score, we have the following:

$$\begin{aligned} S_i &= \sum_k \omega_k^i \sum_{x,y} g_k(x, y) \\ &= \sum_{x,y} \sum_k \omega_k^i g_k(x, y) \end{aligned}$$

In addition, we define M_i as the class activation map for class i , where each spatial element is the product of the weight ω_k^i and the activation $g_k(x, y)$ given by the following:

$$M_i(x, y) = \sum_k \omega_k^i g_k(x, y)$$

Therefore, $S_i = \sum_{x,y} M_i(x, y)$ and $M_i(x, y)$ directly show the importance of the activation at spatial grid (x, y) leading to the classification of an image to class i .

The CAM is a weighted linear sum of the presence of these visual patterns at different spatial locations. By upsampling the class activation map to the size of the input image, we could identify the image regions most relevant to the particular category.

Nodular region segmentation and noise perturbation

We divided each CAM of thyroid nodules into five nearly concentric regions of equal sizes. This was performed first by setting a binarization threshold to 0 to obtain a nodular segmentation with a total area of N . Then, we searched for a second threshold for level setting the CAM to segment the outermost region with the size of $N/5$. We iterated this process until we segmented out the innermost region with the same size.

To illustrate the different importance of each region of the heat map, we added perturbations to the different regions. A commonly employed first-order adversarial attack method—Fast Gradient Sign Method (FGSM) (19)—was used to generate perturbed thyroid ultrasound images.

Let x be the input to the model, y the output to the model (the targets associated with x), ϕ the parameters of a model, and $L(x, y, \phi)$ the cost function used to train the neural network. Then, we can linearize the cost function about the current value of ϕ , acquiring an optimal max-norm constrained perturbation of

$$\delta = \beta \text{sign}(\nabla_x L(x, y, \phi))$$

where β is a predefined perturbation size, which represents the maximum change to pixel values of an image. $\text{sign}()$ is a symbolic function, which is defined as

$$\text{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

In addition, we used the back-propagation to compute the required gradient.

This method can reliably generate perturbations to the input of the model. The single-step FGSM perturbs the original example by a fixed amount along the direction (sign) of the gradient of loss function such that the result from the perturbed image is given by the following:

$$x' = x + \beta \text{sign}(\nabla_x L(x, y, \phi))$$

Statistical methods

To assess the performance of the AI-CADx system, we computed the receiver operating characteristic (ROC) curve and used the area under the curve (AUC) as the evaluation metric. For each nodule, we randomized five times the positions in segmented nodular regions where perturbation noises were applied, and we used the malignancy probabilities averaged over five times of randomized noise perturbations to compute the ROC curves and the subsequent AUC values. To compare individual ROC curves, we performed DeLong's test (22) to evaluate whether the differences in AUC values were statistically significant.

For further statistical significance validation, we evaluated AUC values computed for different regions in each dataset (the division of which is described above), followed by paired t-test for different segmented nodular regions.

Ultrasound feature evaluations in highlighted regions versus the inactivated regions

We randomly selected 100 malignant nodules from our dataset and had the ultrasound features rated by radiologists with 15 years of ultrasound examination experience according to widely used ultrasound-based risk stratification American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) (23) in terms of nodule composition, echogenicity, and echogenic foci. We excluded feature evaluation based on shape and margin attributes, as they, by design, could only be evaluated on the whole rather than on the sub-nodular component levels. We included the statistical evaluation of sub-component localizations for the highlighted and inactivated regions visualized through our proposed malignancy heat map representation, however without any associated risk scores. Each scored feature was weighted against its occurrence frequency and multiplied with the risk points given by ACR TI-RADS criteria to gain an overview of the average ultrasound feature profile of the highlighted regions in contrast to the inactivated regions of our proposed malignancy heat map.

Histopathological correspondence evaluation of malignancy heat map

H&E staining images of postoperative histopathological slides with saved ultrasound images of matched sectional views were reexamined by a senior pathologist with 25 years of work experience with the boundaries and shapes of the malignant regions of thyroid tumors on the H&E slides outlined using software provided by digital pathology slide scanner KF-PRO-005-EX (Konfoong Biotech International Co., Ltd., Yuyao, China).

Results

Heat maps for benign and malignant nodule classifications

Representative examples of the CAMs generated by using AI-SONICTM Thyroid SDK are shown in Figure 1. The first row shows a benign thyroid nodule, and the second row shows a malignant thyroid nodule. The maps highlight the discriminative image regions for thyroid nodule classification. On the left column, zoomed-in thyroid nodule images are shown, whereas the middle column shows the corresponding CAMs. The thyroid images together with their CAMs are superimposed and shown in the original images on the right column. We can see from the exemplars that the malignant nodule has a more complicated CAM profile than the benign one, which may reflect some correlation with the shapes of nodular margins.

Influence of noise perturbations to different CAM regions on thyroid nodule diagnosis

In order to evaluate whether different regions of thyroid nodules may contribute differently to their classifications by the AI-CADx system, we first segmented individual CAMs into five concentric areas of equal sizes according to the method section about nodular CAM segmentation as described above. An illustration of how the segmentation of a nodular CAM looks is given in Figure 2.

We then added gradient sign noise perturbation, which is commonly used for adversarial attacks for CNN models, to each segmented region in the original ultrasound images as shown in Figure 3. The first row shows the noise images added to each segmented region (Figure 2), and the second row shows the noise-perturbed images. The noises added to the original images are barely visible to human eyes but do have a strong impact on the diagnostic performance of the AI-CADx system. We chose the noise magnitude β to be 0.0136 by searching for the maximum absolute gradient of AUC values with respect to the noise magnitudes (Supplementary Figure 1). As a control, the ROC curve and the corresponding AUC value calculated for non-perturbed images are given in Supplementary Figure 2.

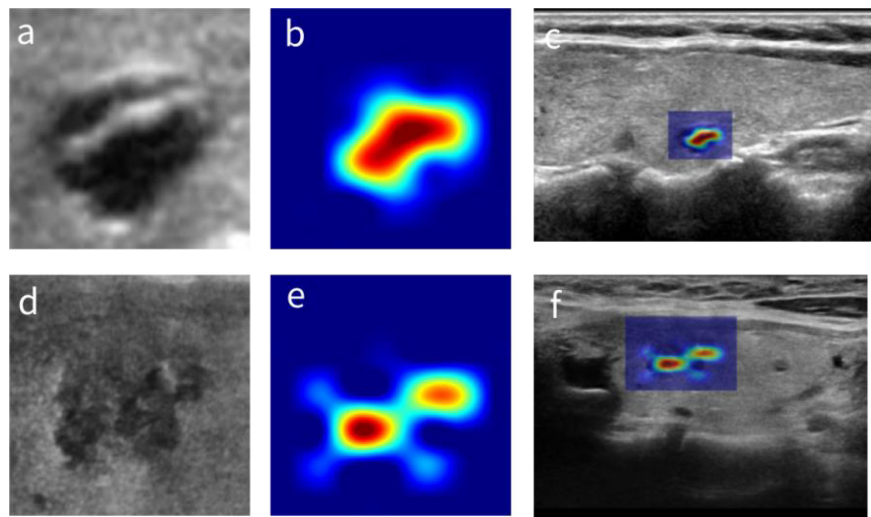


FIGURE 1

The CAMs of thyroid nodules. (A) A benign thyroid nodule. (B) The corresponding CAM. (C) The superposition of panels (A, B) in the original image. (D) A malignant thyroid nodule. (E) The corresponding CAM. (F) The superposition of panels (D, E) in the original image. CAMs, class activation maps.

All thyroid nodule images with noises added to different regions were then classified, and the ROC curves were computed for different noise-perturbed regions, as shown in Figure 3K. To statistically compare the influence of noise perturbation on different regions using ROC curves, we computed pairwise p-values according to DeLong's test (16), with the results shown in Figure 3K. Note that we skipped the statistical comparisons against oneself, as in this case, the p-values were constant 1. It can be seen that it was not the innermost or hottest region (green line for region I) from the CAMs that were worst affected by noises but region IV

(purple line) followed by region III (orange line), suggesting that the tissues surrounding the core area identified by CAMs played a crucial role in the benign and malignant nodule diagnoses.

To further validate this observation, we subdivided the complete dataset into five subsets randomly, summarized in Table 1.

We calculated the AUC values for each dataset in which noise perturbations were applied in the same way as described above to individually segmented nodular regions, and we computed their average values and standard deviations over the five datasets as well

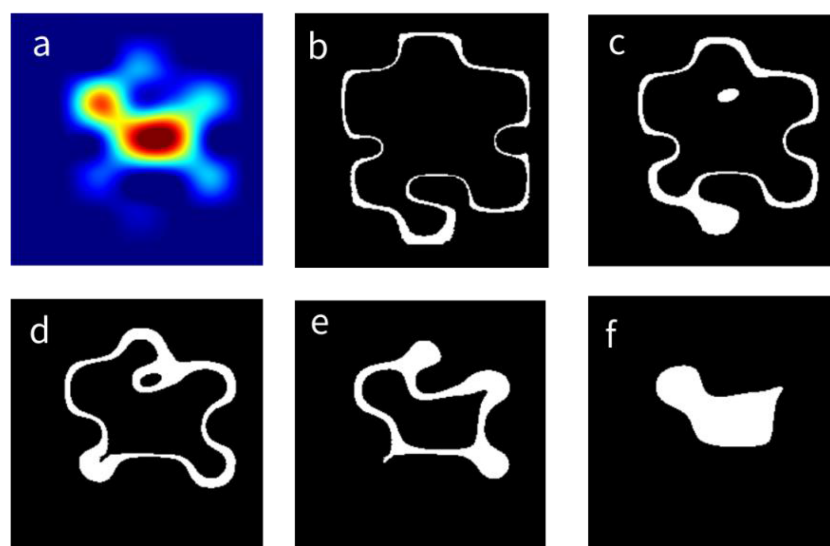
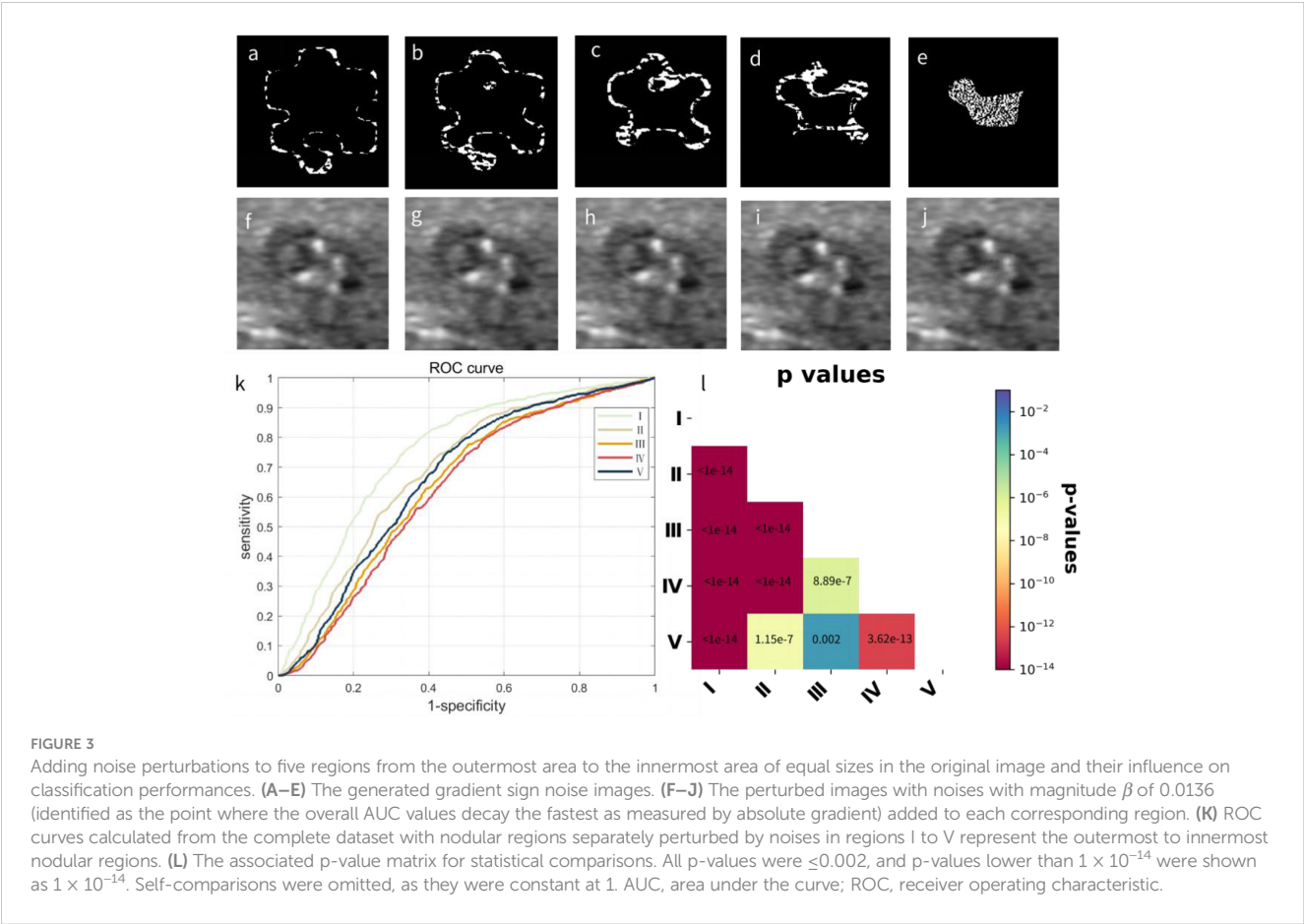


FIGURE 2

Division of a CAM into five equal parts. (A) The CAM of a thyroid nodule. (B–F) The segmented region from the outermost region (labeled as region I) to the innermost region (labeled as region V), respectively. CAM, class activation map.



as the p-value matrix for pairwise comparisons. The corresponding results are given in Figure 4.

Influence of noise perturbations to different malignancy regions on thyroid nodule diagnosis

As cases with suspicious malignant features are more clinically important to be identified in order for subsequent treatment planning, for instance, FNAB, to verify the malignancy status, we modified the conventional CAMs by multiplying the malignancy probability predicted by the AI-CADx model with the CAM specific

for cases suspicious for being malignant to generate malignancy heat maps.

We performed nodular segmentation and applied noise perturbations in the same way as described previously in the Materials and Methods section. In this experiment, we also randomly subset the samples with predicted malignancy scores of higher than 0.4 (below which the probability of being malignant was approximately 3%) to five datasets for the purpose of ensuring reproducibility. The results together with the p-value matrix for pairwise comparisons are summarized in Figure 5. It can be found that in this case, adding noises to innermost region V has the greatest influence on thyroid nodule diagnostic performances, correlating well with the heat intensity profile (Figure 5).

TABLE 1 The subdivided datasets for subsequent nodular region comparison experiment.

Dataset	Total nodules	Benign	Malignant
1	520	320	200
2	520	329	191
3	520	300	220
4	520	322	198
5	522	310	212

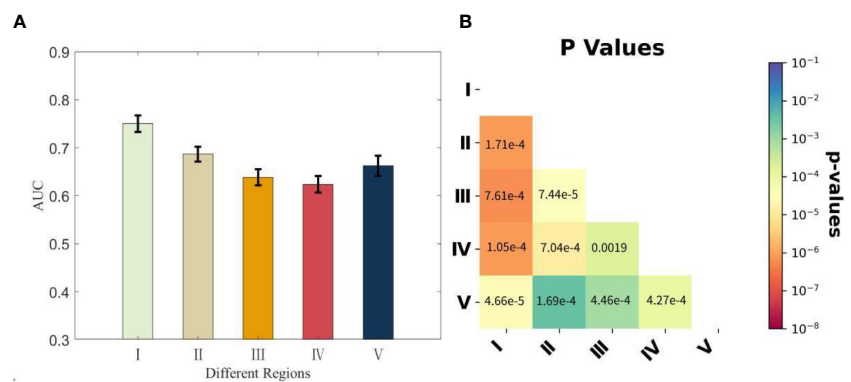


FIGURE 4
The AUC values calculated from the five subdivided datasets with nodular regions separately perturbed by noises and the associated p-value matrix for statistical comparisons. **(A)** Each bar representing the corresponding region is presented with the average AUC values over the five subdivided datasets and the standard deviations. **(B)** All p-values were<0.004, and self-comparisons were omitted, as they were constant at 1. AUC, area under the curve.

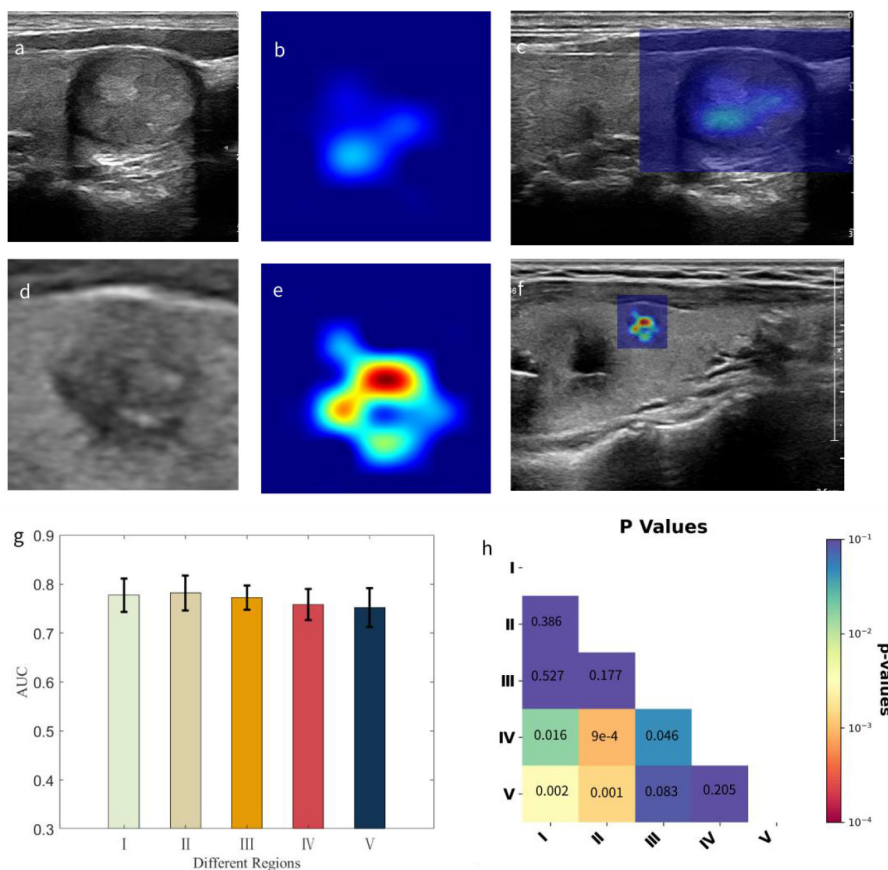


FIGURE 5
Our proposed malignancy heat map and the influence of noise perturbations to five different regions on classification performances. **(A, D)** The cropped zoomed-in images for benign and malignant nodules, respectively. **(B, E)** The corresponding heat maps with color temperature bounded by the predicted malignant probabilities (0.2907 vs. 0.9382) by the AI-CADx system. **(C, F)** The superimposed images for the corresponding nodules and heat maps. **(G)** Each bar representing the corresponding region is presented with the average AUC values over the five subdivided malignant datasets and the standard deviations. **(H)** The associated p-value matrix for statistical comparisons. The smallest p-values were<0.001, and self-comparisons were omitted, as they were constant at 1. AI-CADx, artificial intelligence computer-aided diagnosis; AUC, area under the curve.

Correlation of highlighted regions with higher risks versus inactivated regions of the heat maps

To investigate whether the highlighted regions had a higher correlation with malignancy risks compared to the inactivated regions of our proposed malignancy heat map, radiologists with more than 15 years of ultrasound examination experience evaluated the ultrasound features according to the ACR TI-RADS risk stratification criteria in terms of nodular composition, echogenicity, and echogenic foci, which are applicable to subcomponent evaluations. The results summarized from 100 randomly selected malignant nodules in Table 2 show that the more centrally localized highlighted regions in the malignancy heat maps had higher summed weighted risk scores of 6.04 when compared to the inactivated regions of 4.96, demonstrating a higher correlation with the malignancy risks. On the whole, the highlighted regions turn out to be more hypoechoic and more likely to contain punctate calcifications.

Pathological significance of the heat maps

In order to verify the pathological significance of the heat maps, we first compared the nodule masks generated by the AI system with the masks outlined by the radiologists of different nodule sizes in Figure 6 to show that the basis for computing the heat maps is pathologically relevant. We calculated the dice similarity coefficient (Dice) as a metric to evaluate how well the AI system performs for localizing thyroid nodules within the gland tissues, with the result

shown by the boxplot in Figure 7. It can be found that the masks outputted by the AI system are highly overlapped with the masks delineated by the radiologists (the median dice coefficient >0.9), demonstrating the high accuracy of the AI system in identifying the nodule areas. This verifies that the area of the heat map is mainly in the lesion area, as nodule segmentation is a key step for heat map computations.

To further demonstrate and qualitatively verify the pathological relevance of the computed heat maps, we compared the post-operation pathological images with heat maps computed on ultrasound images shown in Figure 8. It shows that the shapes of cancer cell-rich areas in pathological images outlined by the pathologists have very good correspondence to “hot regions” in our proposed malignancy heat maps, suggesting that the malignancy heat map has promising potential to provide accurate sampling position guidance for FNAB. It is interesting to observe that in Figure 8G where there is clearly one single nodule in the original ultrasound image, the malignancy heat map shows two strongly activated subregions that correspond nicely with the two malignant cell-enriched regions outlined in H&E-stained histopathology images.

Discussion

In this study, we proposed to take advantage of CAMs for differentiating regions in the ultrasound images of thyroid nodules that may contribute differently to the diagnosis by an AI-CADx system. We segmented the CAMs of individual nodule images into five concentric areas of equal sizes, applied adversarial noise

TABLE 2 The ACR TI-RADS feature analysis of CAM highlighted and inactivated regions in 100 malignant nodules randomly selected from the dataset.

Features	CAM hot regions (100)			CAM inactivated regions in the same nodules (100)		
	Frequency	Probability	Weighted score	Frequency	Probability	Weighted score
Localization						
On the margin	27	0.27	–	93	0.93	–
Not on the margin	73	0.73	–	7	0.07	–
Echogenicity						
Very hypoechoic	84	0.84	2.52	43	0.43	1.29
Hypoechoic	16	0.16	0.32	52	0.52	1.04
Isoechoic	0	0.00	0	5	0.05	0.05
Composition						
Solid	100	1.00	2	100	1.00	2
Echogenic foci						
Macro-calcification	0	0.00	0	16	0.16	0.16
Punctate calcification	40	0.40	1.2	14	0.14	0.42
None	60	0.60	0	70	0.70	0
Sum	100	1	6.04	100	1	4.96

Note that the nodular shapes and margin features in ACR TI-RADS are defined for the whole nodules but not sub-nodular regions. Therefore, only relative localizations are provided, which are however not associated with defined risk points.

ACR, American College of Radiology; TI-RADS, Thyroid Imaging, Reporting and Data System; CAM, class activation map.

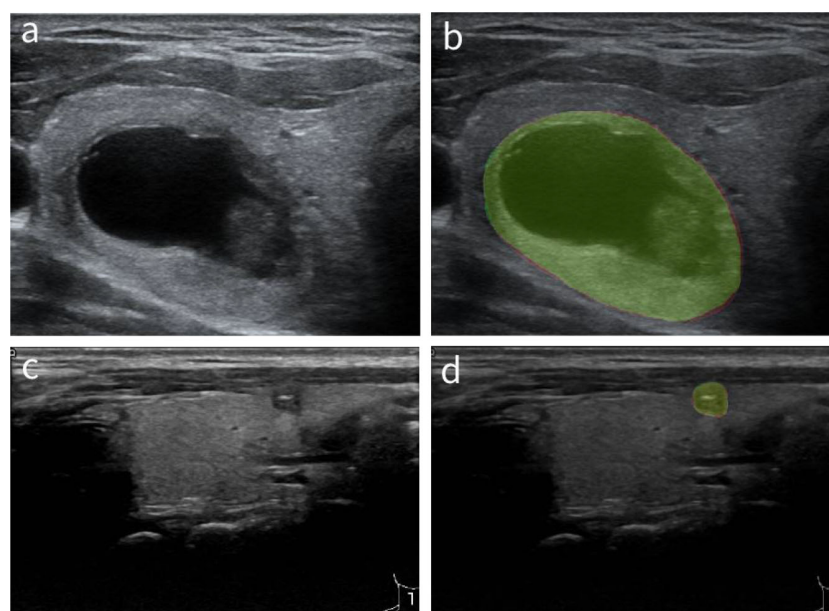


FIGURE 6

Comparison of the masks generated by the AI system and the masks outlined by the radiologists. (A, C) The original thyroid ultrasound images. (B, D) Masks generated by the AI system and radiologists, where the red segments are produced by the AI system, and the green ones are manually drawn by the radiologists. AI, artificial intelligence.

perturbations to the different segmented regions, and evaluated the impact on the diagnostic performances of the AI-CADx system. Our results confirmed that the CAMs can reflect the importance of different degrees in different nodular regions for a CNN-based diagnosis system to make its predictions according to our noise perturbation experiments. Surprisingly or not, it was found that it was not the innermost region or in other words the “hottest” region seen in the CAMs that were most severely influenced by the noise perturbations as one might presumably expect. This phenomenon was again observed in experiments where we randomly divided the original dataset into five subsets and verified by the statistical significance tests, suggesting that the regional sensitivities to noise

perturbations may not perfectly correlate with the heat intensity profile in conventional CAM for nodular malignancy state predictions.

We then tried modifying the conventional CAM by multiplying the heat maps generated by CAM with the predicted malignancy probabilities by the AI-CADx system to produce malignancy heat maps. The resulting heat maps can therefore visually convey information about the malignancy predictions for individual nodules. Furthermore, this malignant nodule-specific modification to the conventional CAM would in principle allow more insightful inspection of regional importance for malignancy assessment and thus provide guidance to doctors about where the FNAB should be mostly directed to. In this case, the innermost nodular area defined by the heat map was found to be the most sensitive region to noise perturbations, which could be due to the multiplication of malignancy probability, which improves its malignancy relevance and subsequently the perturbation responsiveness, though it was not statistically different from the region adjacent to it, which suggests that both regions are the most critical for determining malignant thyroid nodule diagnosis.

By comparing the influences of noises on AI-CADx performance for thyroid nodule diagnosis by perturbing regions segmented according to the conventional CAMs that analyze both benign and malignant nodules and our proposed variant that concentrates on suspicious nodules, it can be seen that noise perturbations to inner regions defined in CAMs had a more dramatic effect on datasets not excluding benign nodules predicted with high confidence. This could be because the inner regions of benign nodules are more vulnerable to noise perturbations, while the malignant nodules are more resistant to noise perturbations. This can also explain why the regional

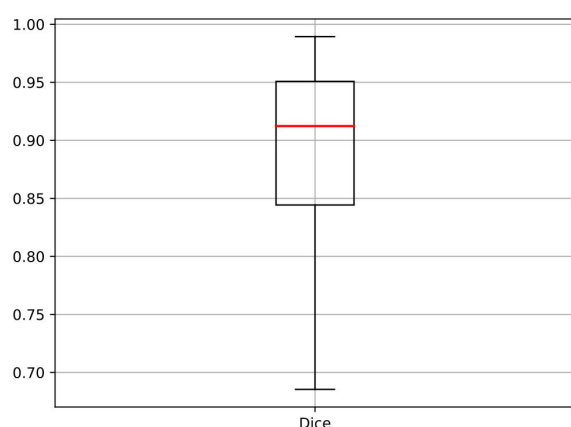


FIGURE 7

The boxplot of the Dice coefficients for quantitative evaluation of the automatically segmented nodules by the AI system in comparison to radiologists' segmentations. AI, artificial intelligence.

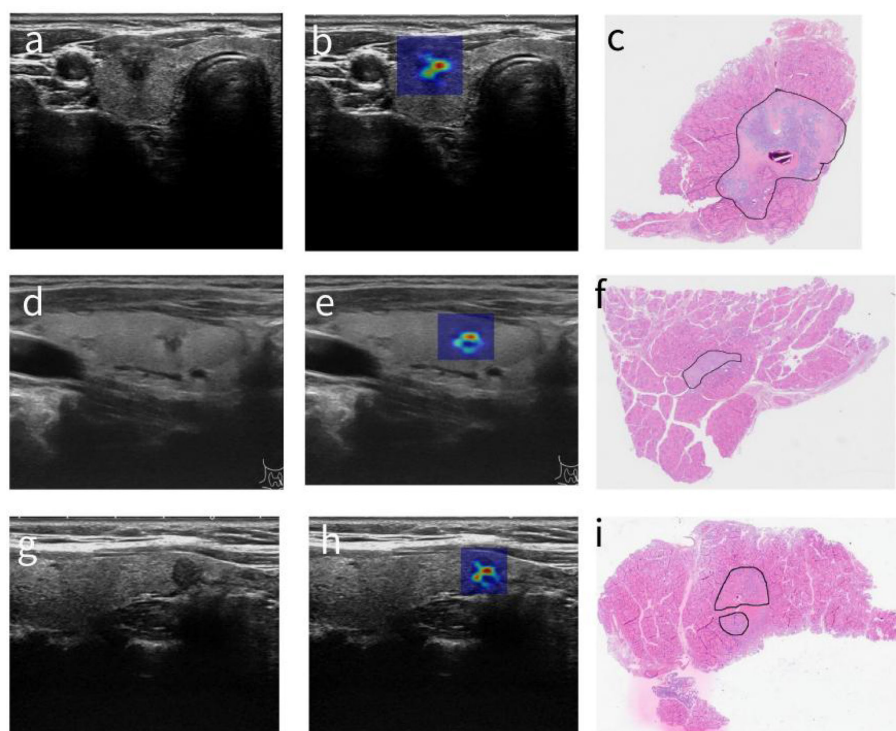


FIGURE 8

Correspondence of post-operation pathological images and malignancy heat maps computed on ultrasound images. (A, D, G) The original thyroid ultrasound images. (B, E, H) The superimposed images for the corresponding nodules and heat maps. (C, F, I) The cropped zoomed-in pathological images.

differences in suspicious malignant nodules were less abrupt. Nonetheless, the central region defined by the malignancy heat map was significantly more sensitive to noise perturbation, suggesting the potential of this technique for differentiating tumor heterogeneity.

As an alternative to the noise perturbation experiments, we have also tried varying binary thresholds to segment the CAMs and evaluated the diagnostic performance of the AI-CADx system directly on the ultrasound images of the regions above the thresholds with the hope of finding an appropriate heat intensity threshold that could permit a satisfactory cutoff to differentiate insignificant regions from regions where special attention would be paid to. However, such a conclusion could not be drawn from this attempt, as the AUC values (Supplementary Figure 3) showed a continual decrease with the heat intensity threshold. This can be attributed to the fact that the marginal shapes of thyroid nodules are crucial for malignancy diagnosis because the segmented regions can have very complicated margins and even hollow structures that can easily confuse the AI-CADx system. Noise perturbation experiments however circumvent this challenge.

As final proof, we qualitatively evaluate the spatial correlation of the malignancy heterogeneity identified by our CAM-based malignancy heat map with the surgical pathology. Due to not only the technical difficulty of automatically registering cross-modality images, i.e., digital histopathology images and ultrasound images, but also more importantly the retrospective nature of this study, we could not specifically collect the ultrasound

images of the very nodular cross-sections that best corresponded to those of the histopathology images. Moreover, arbitrary shape changes can be introduced during the surgical operation and fully automatic registration between digital H&E histopathology images, and ultrasound images would also require training of a sufficient good segmentation model for cancer cell-rich regions in histopathology images, which is currently beyond our capacity of assessable resources. It is however important to note that the ultrasound feature evaluations of the highlighted regions performed by radiologists according to ACR TI-RADS criteria show a higher correlation with malignancy risks compared to the inactivated regions, demonstrating the malignancy relevance of the highlighted regions in our proposed variant of CAM heat map. This can be mainly attributed to two facts: the AI system has high accuracy in localizing thyroid nodules with a median dice coefficient >0.9 , and the AI system has very high diagnostic accuracy with an AUC value of 0.9302 in this study for discriminating between malignant and benign nodules. The nodule segmentation is a crucial step for the computations of heat maps, and it proves that the corresponding area of the heat map is mainly in the lesion area of the nodule. Good performance in distinguishing malignant from benign nodules also plays a vital role in identifying the “hot” regions in the malignancy heat map to be the key regions for the diagnosis of malignant nodules. Furthermore, other recent studies have shown that the AI system has balanced specificity and sensitivity with overall diagnostic accuracy matching high-performing senior radiologists (24), can outperform senior radiologists in diagnosing

rare thyroid carcinomas (25), and can be potentially helpful for discrimination between malignant and benign follicular-patterned thyroid lesions (21). Interestingly, the shapes and subregions of cancer cell-rich pathological images showed good correspondence to the heat maps computed from the ultrasound images, suggesting a promising potential to use the heat map visualization to guide targeted FNAB for more reliable sampling compared with conventional ultrasound-guided sampling.

Ultrasound-guided thyroid FNAB has been shown to improve the sampling accuracy for suspicious nodule identification (26). However, there is a trade-off between the reduction of patient discomfort that would put a constraint on the number of needle passes and the diagnostic accuracy that is limited by specimen adequacy. The CAM-based heat maps computed from AI-CADx systems with diagnostic performances comparable to or even better than those of senior radiologists (27–29) on ultrasound images with the capability of differentiating regional importance for malignancy diagnosis may provide additional guidance to localize diagnosis-enabling nodular regions, especially large ones, for more accurate sampling, given that the number of needle passes has to be limited. In addition, FNAB-based cytopathological examination is acknowledged to have a limitation in diagnosing follicular-patterned thyroid lesions (FPTLs) (30–32), while the AI system was found to be 69% accurate in differentiating thyroid follicular carcinoma from benign FPTL cases (21), suggesting that the heat maps developed on top of the AI system may provide better guidance than plain ultrasound for FPTL sampling by FNAB to help with newly developed proteomics-based diagnosis (33). Of course, for smaller thyroid nodules, it may be difficult to precisely guide FNAB of the segmented nodular regions based on the malignancy heat map. Meanwhile, it must be noticed that the FNAB selection of thyroid nodules has become increasingly conservative in clinical practice. FNAB for malignant suspicious thyroid nodules recommended by guidelines (23, 34) is commonly performed for nodules with the smallest diameter ≥ 10 mm. Therefore, it is of clinical significance for this study to guide FNAB of relatively large thyroid nodules with the heat maps. It may also be of special interest to investigate whether CAM-based heat maps on ultrasound images can be helpful for guiding core needle biopsies of, for instance, liver lesions, which was shown to have a complication rate of 10.6% for repeated biopsies with a diagnostic accuracy of 83.3% (35).

Furthermore, currently, the generation of CAMs is based on an AI-CADx system trained on static images. For real-world clinical applications, it will be beneficial to have such heat maps dynamically generated in real-time during ultrasound scanning, which would require the corresponding AI-CADx system to be able to operate in a dynamic mode with high diagnostic accuracy. Practically speaking, ultrasound reflections from needles might interfere with accurate heat map visualization in real time. Another noteworthy limitation of the proposed CAM and its variant presented in this study is that their visualization is currently limited to being two-dimensional and thus not very suited yet to visualizing the degree of malignancy suspiciousness

of a targeted cross-sectional plane relative to that of the planes above and below. If a three-dimensional heat map visualization is to be developed such that additional guidance about how deep the needles shall be inserted into the nodules to acquire samplings can be possible. In addition to the applied basic CAM method, there are also other published variants such as Grad-CAM (36), Score-CAM (37), and Ablation-CAM (38). These methods of generating heat maps can in principle be investigated as well. However, it is for future studies to evaluate which heat map generation techniques can be the most useful for guiding FNAB. This work is mainly to show that CAM-based heat maps can visualize intra-nodular malignancy heterogeneity, and this may recommend better sites for FNAB sampling, which is to be evaluated in a separate study.

Conclusion

The CAM and its variant generated on ultrasound images through a highly accurate AI-CADx system can provide differential importance of nodular regions for tumor malignancy prediction, which was validated by adding noise perturbations to different regions of thyroid nodules. Our proposed malignancy heat map offers quantitative visualization of malignancy heterogeneity within a tumor, and the highlighted regions are better correlated with the malignancy risk than the inactivated regions. The good spatial correspondence with post-operation pathology warrants clinical interests to investigate further whether such AI-based malignancy-heterogeneity visualization techniques can provide targeted guidance for needle-based aspiration biopsies of tumors in comparison with plain ultrasound imaging to improve sampling accuracy and reduce complications that may associate with the procedures.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: It is not within the agreement with the participants to make the original dataset publicly available. The data that support the findings of this study are available on reasonable request from corresponding author L.W. after formal approval by the concerned Chinese regulating authorities.. Requests to access these datasets should be directed to wanglj844@zjcc.org.cn.

Ethics statement

The studies involving human participants were reviewed and approved by The Cancer Hospital of the University of Chinese Academy of Sciences but waived in view of the retrospective nature of the study and all involved ultrasound image reviews being performed were part of the clinical routine. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization: LPW and LX. Methodology: YW, JY, and WL. Software: YW, WL, and JY. Validation: LPW, LJW, and LX. Formal analysis: YW and DX. Investigation: LPW, YW, DX, and LJW. Resources: LJW and LX. Data curation: LPW, LJW, and DX. Writing—original draft preparation: YW and LX. Writing—review and editing: LX. Supervision: LJW and LX. Project administration: LJW and LX. Funding acquisition: DX and LPW. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (No. 82071946), the Zhejiang Provincial Natural Science Foundation of China (No. LY20H180001, LZ21F030001), Project of Zhejiang Medical and Health Science and Technology Plan (No. 2022KY661, 2021KY099, 2022KY110), and the University Cancer Foundation via the Sister Institution Network Fund at the University of Texas MD Anderson Cancer Center.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest* (2009) 39(8):699–706. doi: 10.1111/j.1365-2362.2009.02162.x
- Gharib H, Goellner JR. Fine-needle aspiration biopsy of the thyroid: an appraisal. *Ann Intern Med* (1993) 118(4):282–9. doi: 10.7326/0003-4819-118-4-199302150-00007
- Amrikachi M, Ramzy I, Rubinfeld S, Wheeler TM. Accuracy of fine-needle aspiration of thyroid: a review of 6226 cases and correlation with surgical or clinical outcome. *Arch Pathol Lab Med* (2001) 125(4):484–8. doi: 10.1089/108729001750171399
- Sung JY, Na DG, Kim KS, Yoo H, Lee H, Kim JH, et al. Diagnostic accuracy of fine-needle aspiration versus core-needle biopsy for the diagnosis of thyroid malignancy in a clinical cohort. *Eur Radiol* (2012) 22(7):1564–72. doi: 10.1007/s00330-012-2405-6
- Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. *Nature* (2013) 501(7467):328–37. doi: 10.1038/nature12624
- Chmielik E, Rusinek D, Oczko-Wojciechowska M, Jarzab M, Krajewska J, Czarniecka A, et al. Heterogeneity of Thyroid Cance. *Pathobiology* (2018) 5(1-2):117–29. doi: 10.1159/000486422
- Lee YH, Baek JH, Jung SL, JY K, JH K, Shin JH. Ultrasound-guided fine needle aspiration of thyroid nodules: a consensus statement by the Korean society of thyroid radiology. *Korean J Radiol* (2015) 16(2):391–401. doi: 10.3348/kjr.2015.16.2.391
- Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs L, et al. American Association of clinical endocrinologists, American college of endocrinology, and associazione Medici endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules—2016 update. *Endocr Pract* (2016) 22(5):622–39. doi: 10.4158/EP161208.GL
- Jung CK, Baek JH, Na DG, Oh YL, Yi KH, Kang HC. 2019 Practice guidelines for thyroid core needle biopsy: a report of the clinical practice guidelines development committee of the Korean thyroid association. *J Pathol Transl Med* (2020) 54(1):64–86. doi: 10.4132/jptm.2019.12.04
- Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer* (2018) 18(8):500–10. doi: 10.1038/s41568-018-0016-5
- Ma JL, Wu F, Jiang T, Zhu J, Kong KX. Cascade convolutional neural networks for automatic detection of thyroid nodules in ultrasound images. *Med Phys* (2017) 44(5):1678–91. doi: 10.1002/mp.12134
- Ma JL, Wu F, Zhu J, Xu D, Kong KX. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics* (2017) 73:221–30. doi: 10.1016/j.ultras.2016.09.011
- Jia X, Ma Z, Kong D, Li Y, Hu H, Guan L, et al. Novel human artificial intelligence hybrid framework pinpoints thyroid nodule malignancy and identifies overlooked second-order ultrasonographic features. *Cancers* (2022) 14:4440. doi: 10.3390/cancers14184440
- Dai W, Cui Y, Wang P, Wu H, Zhang L, Bian Y, et al. Classification regularized dimensionality reduction improves ultrasound thyroid nodule diagnostic accuracy and inter-observer consistency. *Comput Biol Med* (2023) 154:106536. doi: 10.1016/j.combiomed.2023.106536
- Zhou B, Khosla A. Learning deep features for discriminative localization, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (CVPR)* (2016) pp. 2921–9.
- Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. *Comput Sci - Comput Vision Pattern Recognition* (2017). doi: 10.48550/arXiv.1711.05225
- Chan HP, Hadjiiski LM, Samala RK. Computer-aided diagnosis in the era of deep learning. *Med Phys* (2020) 47(5):e218–27. doi: 10.1002/mp.13764

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1136922/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Influence of adding gradient sign noise perturbations to ultrasound images of different magnitudes b. (a and d) The original image of thyroid nodule and its surroundings. (b and e) The generated gradient sign noises with the respective magnitude of 0.0136 and 0.05. (c and f) The resulting images corresponding to noise perturbations given by panels b and e. (g) The measurement of how the diagnosis performances (by AUC values) of the AICADx system are influenced by noise perturbations with magnitudes b ranging from 0 to 0.1.

SUPPLEMENTARY FIGURE 2

The ROC curve of the AI-CADx system for thyroid nodule diagnosis and the corresponding AUC value.

SUPPLEMENTARY FIGURE 3

The diagnostic performance of the AI-CADx system directly on segmented regions above the binarization CAM threshold (represented as a ratio to the maximum temperature intensity 255).

18. Wang LT, Zhang L, Zhu MJ, Qi XF, Yi Z. Automatic diagnosis for thyroid nodules in ultrasound images by deep neural networks. *Med Image Anal* (2020) 61:101665. doi: 10.1016/j.media.2020.101665
19. Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. *Comput Sci* (2014). doi: 10.48550/arXiv.1412.6572
20. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. *Internatl Conference Machine Learning* (2019) 97:6105–14. doi: 10.48550/arXiv.1905.11946
21. Xu D, Wang Y, Wu H, Lu W, Chang W, Yao J, et al. An artificial intelligence ultrasound system's ability to distinguish benign from malignant follicular-patterned lesions. *Front Endocrinol (Lausanne)* (2022) 13:981403. doi: 10.3389/fendo.2022.981403
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* (1988) 44(3):837–45. doi: 10.2307/2531595
23. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *J Am Coll Radiol* (2017) 14(5):587–95. doi: 10.1016/j.jacr.2017.01.046
24. Guo F, Chang W, Zhao J, Xu L, Zheng X, Guo J. Assessment of the statistical optimization strategies and clinical evaluation of an artificial intelligence-based automated diagnostic system for thyroid nodule screening. *Quant Imaging Med Surg* (2023) 13(2):695–706. doi: 10.21037/qims-22-85
25. Wang Y, Xu L, Lu W, Kong X, Shi K, Wang L, et al. Clinical evaluation of malignancy diagnosis of rare thyroid carcinomas by an artificial intelligent automatic diagnosis system. *Endocrine* (2022) 80(1):93–9. doi: 10.1007/s12020-022-03269-4
26. Tsao GJ, Orloff LA. Clinician-performed thyroid ultrasound-guided fine-needle aspiration. *Otolaryngol Clin North Am* (2014) 47(4):509–18. doi: 10.1016/j.otc.2014.04.005
27. Buda M, Wildman-Tobriner B, Hoang JK, Thayer D, Tessler FN, Middleton WD, et al. Management of thyroid nodules seen on US images: deep learning may match performance of radiologists. *Radiology* (2019) 292(3):695–701. doi: 10.1148/radiol.2019181343
28. Peng S, Liu YH, Lv WM, Liu LZ, Zhou Q, Yang H, et al. Deep learning-based artificial intelligence model to assist thyroid nodule diagnosis and management: a multicentre diagnostic study. *Lancet Digit Health* (2021) 3(4):e250–9. doi: 10.1016/S2589-7500(21)00041-8
29. Mai W, Zhou M, Li J, Yi W, Li S, Hu Y, et al. The value of the demetics ultrasound-assisted diagnosis system in the differential diagnosis of benign from malignant thyroid nodules and analysis of the influencing factors. *Eur Radiol* (2021) 31(10):7936–44. doi: 10.1007/s00330-021-07884-z
30. Poller DN, Ibrahim AK, Cummings MH, Mikel JJ, Boote D, Perry M. Fine-needle aspiration of the thyroid. *Cancer* (2000) 90(4):239–44. doi: 10.1002/1097-0142(20000825)90:4<239::aid-cnrcr7>3.0.co;2-s
31. Busseniers AE, Oertel YC. Cellular adenomatoid nodules" of the thyroid: review of 219 fine-needle aspirates. *Diagn Cytopathol* (1993) 9(5):581–9. doi: 10.1002/dc.2840090523
32. Rossi ED, Adeniran AJ, Faquin W. Pitfalls in thyroid cytopathology. *Surg Pathol Clin* (2019) 12(4):865–81. doi: 10.1016/j.path.2019.08.001
33. Sun Y, Selvarajan S, Zang Z, Liu W, Zhu Y, Zhang H, et al. Artificial intelligence defines protein-based classification of thyroid nodules. *Cell Discovery* (2022) 8(1):85. doi: 10.1038/s41421-022-00442-x
34. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* (2016) 26(1):1–133. doi: 10.1089/thy.2015.0020
35. Kim JS, Won HJ, Lee SJ, Kim SY, Shin YM, Kim PN. Utility and safety of repeated ultrasound-guided core needle biopsy of focal liver masses. *J Ultrasound Med* (2018) 37(2):447–52. doi: 10.1002/jum.14358
36. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vision* (2020) 128(2):336–59. doi: 10.1007/s11263-019-01228-7
37. Wang H, Wang Z, Du M, Yang F, Zhang Z, Ding S, et al. Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA (2020) 111–119. doi: 10.1109/CVPRW50498.2020.00020
38. Desai S, Ramaswamy HG. Ablation-cam: visual explanations for deep convolutional network via gradient-free localization. 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA (2020) 972–980. doi: 10.1109/WACV45572.2020.9093360



OPEN ACCESS

EDITED BY

Alla Reznik,
Lakehead University, Canada

REVIEWED BY

Abdelbaset Mohamed Elsbali,
Al Jouf University, Saudi Arabia
Zuheir Alshehbi,
Tishreen University, Syria

*CORRESPONDENCE

Marina Sekacheva

✉ Sekacheva_m_i@staff.sechenov.ru

RECEIVED 02 August 2022

ACCEPTED 17 April 2023

PUBLISHED 03 May 2023

CITATION

Sekacheva M, Boroda A, Fatyanova A,
Rozhkov A and Bagmet N (2023) Clinical
validation of the novel CLIA-CA-
62 assay efficacy for early-stage
breast cancer detection.
Front. Oncol. 13:1009863.
doi: 10.3389/fonc.2023.1009863

COPYRIGHT

© 2023 Sekacheva, Boroda, Fatyanova,
Rozhkov and Bagmet. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Clinical validation of the novel CLIA-CA-62 assay efficacy for early-stage breast cancer detection

Marina Sekacheva^{1*}, Alexander Boroda¹, Anastasia Fatyanova¹,
Alexander Rozhkov¹ and Nikolai Bagmet²

¹World-Class Research Center "Digital Biodesign and Personalized Healthcare", Sechenov First Moscow State Medical University, Moscow, Russia, ²Department of Biliary, Hepatic, and Pancreatic Surgery, B.V. Petrovsky Russian Scientific Center of Surgery, Moscow, Russia

Background: Without organized screening programs up to 60-70% of breast cancers are diagnosed at advanced stages that have significantly lower five-year survival rate and poorer outcomes, which is a serious global public health problem. The purpose of the blind clinical study was the assessment of the novel *in-vitro* diagnostic chemiluminescent CLIA-CA-62 assay for early-stage breast cancer detection.

Methods: Blind serum samples of 196 BC patients with known TNM staging, 85% with DCIS, Stage I & IIA, and 73 healthy control subjects were analyzed with the CLIA-CA-62 and CA 15-3 ELISA assays. Results were also compared to the pathology findings and to published data from mammography, MRI, ultrasound, and multi-cancer early detection test (MCED).

Results: The CLIA-CA-62 overall sensitivity for BC was 92% (100% for DCIS) at 93% specificity and it decreased in invasive stages (Stage I=97%, Stage II=85% and Stage III=83%). For the CA 15-3 assay sensitivity was 27-46% at 80% specificity. Sensitivity for mammography was 63-80% at 60% specificity, depending on the stage and the parenchymal density.

Conclusion: These results demonstrate that CLIA-CA-62 immunoassay could prove useful as a supplement to current mammography screening and other imaging methods, thus increasing the diagnostic sensitivity in DCIS and Stage I breast cancer detection.

KEYWORDS

CLIA, tumor, CA-62, CA 15-3, breast cancer, ductal carcinoma *in-situ* (DCIS), screening, mammography

1 Introduction

According to current global cancer statistics, female breast cancer has become the most commonly diagnosed cancer globally (more than 2 million estimated cases in 2022 worldwide) surpassing lung cancer (1). The global annual percent change (APC) for BC mortality increased by 0.23% (2). The global deaths from breast cancer increased by 83.95% since 1990 (95% UI: 70.07–96.74%), with 685 000 new cases in 2022 (1). Asia takes a leading place for cancer incidence, followed by Europe and North and South America. For many low-and-middle-income countries from 40% to 70% of breast cancer is detected at advanced stages associated with lower five-year survival rates, which represents a serious global health problem (3). The target population for breast cancer awareness are women over 40 years of age, since breast cancer incidence rate increases with age, hereditary breast, and ovarian cancer syndrome, BRCA2 – germline mutations, and female sex hormones fluctuations from 60 cases per 100,000 in women 30–40 years of age to an average of 430 cases per 100,000 in women 65–75 years of age (4).

The following methods are being used for breast cancer diagnostics, such as bilateral and digital mammography, a conventional ultrasound of the mammary glands for women under 40 years of age, ultrasound elastography, and MRI scanning with contrast. Regardless of all the drawbacks of breast cancer mammography screening of women, such as missing some advanced cancers and producing 3/4 of “suspicious mammograms” associated with benign breast diseases, - mammography continues to be the only proven screening test to decrease mortality from breast cancer (5). At present, only a few countries with a high development index have improved prevention measures and implemented population-oriented breast cancer screening programs as well as improved quality of cancer care, allowing the detection of ~80% of early-stage breast cancer (1). The biggest obstacle to the overall screening approach worldwide is the very high cost of the organized screening programs, which have significant budget implications, depending on the size of the population and the healthcare system resources involved. For instance, the overall costs for annual breast cancer screening of 1,000 women in the general population of Canada are approximately \$16.0 million as a lifetime expense (ages 40 to 74) (6). As opposed to the North America and some Western European countries practice, the majority of countries worldwide, including Asia, Arabic, and some Eastern European countries use mammographic screening as an opportunistic diagnostic method since there are no population-based mammographic screening programs (7). To resolve the growing breast cancer problem, it is critical to find an economically viable solution for the prevention of advanced breast malignant disease through early breast cancer detection and optimal access to treatment.

According to the tumor, node, metastasis (TNM) staging classification based on AJCC 8th edition Early-stage breast cancer refers to a malignant neoplasm that has not spread beyond the breast or the axillary lymph nodes (8). This includes Tis – Ductal carcinoma *in situ* or Paget's disease of the breast with no associated tumor mass, Stage I (T1aN0M0, T1bN0M0, and T1cN0M0), Stage

IIA (T0N1M0, T1N1M0), and Stage IIB (T2N0M0, T2N1M0, T3N0M0).

For decades cancer biomarkers were extensively used to detect, diagnose, or manage certain types of cancer within the standard of care in many parts of the world. These biomarkers include different types of glycoproteins or various genes known to be associated with cancer, which are formed within the growth of a neoplasm (9). The detection and identification of such cancer biomarkers in a patient's body fluids provide valuable data in regard to the diagnosis of invasive breast cancers, providing prognostic information and predicting response to a chosen therapy, and a selection of the strategy for treating cancer, which leads to improved outcomes. However, there is a limitation of using circulating biomarkers related to their low sensitivity (10–30%) for detecting early stages of breast cancer (10).

In recent years, our cancer research group has been studying various combinations of biomarkers (CA-125, CA 15-3, CEA, CYFRA 21.1, D-dimer, HE4 etc.) for their effective use in screening for different types of cancers (11, 12), including breast cancer (13). In this regard, we aimed to identify novel emerging biomarkers for their successful use in combination with other well-known cancer markers, which can significantly improve the accuracy of cancer screening using a classification model. Thus, in the last 5 years, our attention has been drawn to various pilot clinical studies that were carried out by our colleagues using a novel marker for epithelial carcinomas CA-62 and the results were presented at various international symposiums.

The purpose of this study was the assessment of the diagnostic characteristics of the novel FSSH-approved (Federal Service for Surveillance in Healthcare of the Russian Federation) *in vitro* diagnostic chemiluminescent immunoassay (IVD CLIA-CA-62) for early-stage breast cancer detection as compared to healthy controls. The same samples were analyzed with FSSH-CA 15-3 ELISA assay, and the results were compared to histopathologic diagnosis used as a gold standard.

The significance of this study is related to the unique qualities of the highly sensitive CA-62 marker, which allows detecting it in the blood of patients with ductal carcinoma *in situ* (DCIS) and Stage I breast cancer present in asymptomatic women. This study has the potential to provide insight into the usefulness of the CA-62 biomarker as a first-line test to select subjects at high risk for developing breast cancer (BC) who need further mammography, potentially avoiding radiological exposure in low-risk BC individuals who test negative.

CA 15-3 ELISA-BEST assay is based on the well-known cancer marker CA 15-3, an O-glycoprotein member of the mucin family commonly used for breast cancer control management (11, 14). It is a protein product of the MUC-1 gene, which is shed into the bloodstream from adenocarcinomas in a reduced glycosylated form. Despite its low sensitivity for early stages of breast cancer detection, CA 15-3 cancer antigen is extensively used for cancer treatment monitoring in combination with clinical examination and various imaging methods and for early detection of cancer recurrence (14, 15).

Human CLIA-CA-62 immunoassay is based on the novel marker for epithelial carcinomas CA-62, which is a carcinoma-

specific mesenchymal marker, expressed on the epithelial cell surface of the EMT-transformed undifferentiated stem cells from the onset of cancer development. There are some previous publications describing CA-62, a patented set of reagents CLIA-CA-62 based on the biomarker CA-62 for early cancer detection (16–18) and monitoring response to chemotherapy (19). A marker for epithelial carcinomas CA-62 represents a family of low-weight membrane transport N-glycoproteins that bind alpha-fetoprotein (AFP) using a special combination of the branched polysaccharides, which are located on the mesenchymal cells' surface and function by Clathrin-mediated endocytosis. Tumor cells release into the blood two main soluble cytoplasmic fractions of N-glycoprotein that are detected by the antibody used in the test. This allows to quantitatively measure a serum level of CA-62 antigen using a specific chemiluminescent assay CLIA-CA-62 intended to help with the medical decision-making process and recommended for early cancer detection in combination with clinical data and other diagnostic procedures (16, 17).

Test performance of both assays was also compared to histopathological findings and to published data for mammography, conventional ultrasound, ultrasound elastography, MRI, and also to blood-based multi-cancer early detection test MCED (from GRAIL Inc.).

2 Materials and methods

2.1 Study design and participants

2.1.1 Patients

Patients with histopathology-confirmed breast cancer before the treatment (N=57) and healthy control subjects (N=73) were enrolled in an observational clinical study in 2018 at the Institute for Personalized Medicine of the Sechenov First Moscow State Medical University ("Sechenov University"), Moscow, Russia. Inclusion criteria for breast cancer patients: women of any race and ethnicity between the ages of 25 and 80, who have been diagnosed with either DCIS or Stage I, Stage IIA-B, or Stage III of primary breast tumor with or without lymph node metastasis; histopathological confirmation of breast cancer, which was used for definitive diagnosis of the breast disease. Exclusion criteria for this study included factors such as age before 25 and above 80, more advanced stage of disease (Stage IV), and previous treatment history.

Breast tissues were collected from resected breast tumors at the time of mastectomy or lumpectomy, fixed in formaldehyde, and embedded in paraffin. The tissues were cut into sections and stained with hematoxylin/eosin. The diagnosis of a benign or malignant breast tumor was confirmed by certified pathologists. Histopathological classification and staging were performed according to AJCC eighth edition (8).

Healthy control subjects were selected based from a large pool of apparently healthy individuals on matching variables of interest such as age (from 25 – 80), same gender, any race and ethnicity, free of cancer, with normal biochemical and full blood count reference intervals seen in a healthy reference population according to the international standard ISO 15189:2012. Exclusion criteria for healthy control subjects included several factors, such as another

gender (men), age before 25 and after 80, presence of comorbidities or verified breast benign disease or breast cancer.

2.1.2 Ethics approval and consent to participate

The study was approved by the Local Ethics Committee of Sechenov First Moscow State Medical University. All patients were given informed consent to participate in the study. In total 57 patients with histopathology-confirmed BC and 73 healthy control subjects were included in the analysis. Serum samples were collected at the Sechenov University Hospital after overnight fasting and delivered to the Clinical laboratory.

2.1.3 Serum samples

The total of 269 blinded serum samples included 196 breast cancer patients with known TNM classification (8, 20) and 73 healthy control subjects. Sera from healthy control subjects (N=73) and pre-treatment breast cancer sera (N=57) were collected at the Sechenov University Hospital after overnight fasting and delivered to the Clinical laboratory, processed and stored at –86°C until they were analyzed for CA-62 and CA 15-3 markers. Another set of archived histopathology-confirmed breast cancer sera (N=139) was obtained from the Biospecimen bank ProMedDX LLC, MA, USA. The entire set of serum samples was separated by centrifugation (1300 g, 10 minutes) in BD SST tubes with silica clot activator, and separating polymer gel, heat-inactivated at 56°C for 30 min using standard operating procedures for serum collection (21, 22), and stored at –86°C until used. Serum samples were collected under an IRB-approved protocol from Federal licensed/registered facility following GMPs. The majority (85%) of cancer samples were from patients with Stage I and Stage II cancers, as well as DCIS (T0). The baseline characteristics of the studied serum samples are presented in Table 1.

2.1.4 Study design

The entire blind set of serum samples (N=269) was analyzed using an FSSH-approved IVD medical device CLIA-CA-62 based on a competitive chemiluminescent assay and 177 BC cases together with 73 healthy control subjects were tested by another FSSH-approved IVD medical device ELISA-CA15-3 based on a sandwich enzyme-based immunoassay. A controlled blind clinical study using serum samples from histopathologically confirmed patients was carried out at Sechenov University, Moscow. A clinical study methodology is presented in Figure 1. The blinding was carried out by experts from the external independent laboratory of the Federal Service for Surveillance in Healthcare of the Russian Federation. Such study design provides a high level of internal validity and allows avoiding any bias, chance or confusion.

2.2 Human CLIA-CA-62 immunoassay

Measurements of CA-62 cancer antigen in patients' sera were performed using the *in vitro* diagnostic immunoassay from JVS Diagnostics LLC, Moscow, Russia (Lot# CLIA-CA-62-200221), according to the manufacturer's instruction. IVD CLIA-CA-62 chemiluminescent immunoassay approved by the Federal Service

TABLE 1 Baseline characteristics of the analyzed serum samples.

	Breast cancer (N = 196)	Normal control subjects (N = 73)	Total (N = 269)
Age, years			
Mean (SD)	61years	55 years	61 years
Median	63years	54 years	62 years
Range	25 – 93 years	45 –70 years	25 – 93 years
Age group, n (%)			
< 50 years	31 (15.8%)	21 (28.8%)	52 (19.3%)
50 – 60 years	36 (18.4%)	35 (47.9%)	71 (26.4%)
61 – 65 years	34 (17.4%)	9 (12.3%)	43 (16.0%)
>65 – 69 years	32 (16.3%)	6 (8.2%)	38 (14.1%)
>70 – 79 years	49 (25.0%)	2 (2.7%)	51 (19.0%)
>80 years	14 (7.1%)	0 (0%)	14 (5.2%)
Race, n (%)			
Caucasian	167 (85.2%)	73 (100%)	240 (89.2%)
Asian	29 (14.8%)	0 (0%)	29 (10.8%)
Clinical cancer stage, with TNM classification n (%)			
Stage 0 (DCIS, pT0N0M0)	14 (7.1%)	none	
Stage I A, B- (pT1N0M0)	98 (50.0%)		
Stage IIA - (pT1-2N0-1M0,	56 (28.6%)		
Stage II B - (pT2-3N0-1M0	17 (8.67%)		
Stage III A – (pT2-3N1M0)	4 (2.04%)		
Stage III B – (pT1-4N2M0)	7 (3.57%)		
Stage IV – (pT0-NxM1)	0 (0%)		
Region, n (%)			
USA	139 (70.9%)	0 (0%)	139 (51.7%)
Russian Federation	57 (29%)	73 (100%)	130 (49.3%)
Method of cancer diagnosis, n (%)			
Identified by screening test	139 (70.9%)	0 (0%)	139 (51.7%)
Identified by clinical presentation	57 (29%)	73 (100%)	130 (49.3%)

for Surveillance in Healthcare of the Russian Federation (FSSH) is intended for the quantitative measurement of carcinoma-specific antigen CA-62 in human blood serum. A set of reagents CLIA-CA-62 is a one-step solid-phase competitive chemiluminescent immunoassay, in which a competition takes place between the carcinoma-specific antigen CA-62 contained in the test sample and the labeled cancer antigen CA-62-Acrininium (Acrininium NHS-ester) for binding to monoclonal antibodies (Mabs) to glycoprotein CA-62, immobilized on the solid phase (96-well plate) (18). The human CLIA-CA-62 test kit contains all the required sets of reagents to analyze 48 samples in duplicates, including the CA-62 standard calibrators. During 2-hour incubation, a solution containing a diluted serum sample (1:5), 50 µl of Positive control sample, 50 µl of Standard Calibrators CA-62, and 50 µl of the labeled cancer antigen 400 ng/ml solution were added to the wells with immobilized antibodies, thereafter the sorbent was washed away from unbound components. A series of standard calibrators CA-62 were tested simultaneously with the specimens to plot a Logit-Log calibration curve for the quantitative determination of the glycoprotein CA-62 in measurement units (U/ml) of the unknown samples. The measurable luminescent signals from the immune complexes {anti-CA62-Mab-(CA-62-Acrininium ester)}

formed on the solid phase were then recorded immediately after the injection of the activating buffer solution (integration time 0.3 sec) in the wells, which induces the photon emission, detected by a flash chemiluminescence reader. The intensity of the luminescent signal is inversely proportional to the concentration of the measured analyte in the specimen. Samples with CA-62 units higher than the largest CA-62 calibrator were diluted accordingly with a working buffer followed by the determination of the exact concentration by multiplying on the dilution factor.

Measurements were made following the manufacturer's instructions. Sensitivity of the assay: 35 U/ml; the assays had linearity of 91 and 105% over the measurement range of 1250 -10000 U/ml. The total analysis time was 4 hours. Detection method: flash chemiluminescence using Tecan Spark (Tecan Trading AG, Switzerland, EU). The cut-off value (5000 U/ml) used was recommended by the manufacturer based on CA-62 levels in sera from 353 healthy individuals 18 – 65 years old. All the test samples were done in duplicates using Standard calibrators CA-62 and Positive control samples as reference standards included in the set of reagents. 95% of the sera from healthy control subjects were above the limit of detection (LOD). The intra-assay coefficient of variation was ≤ 10%, over the range of concentrations.

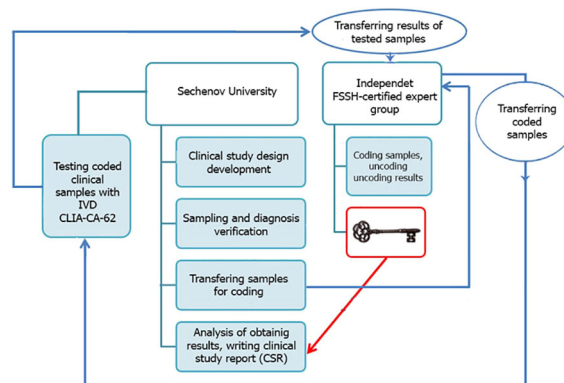


FIGURE 1
A clinical study methodology.

2.3 The sandwich CA 15-3 ELISA

Measurements of CA 15-3 cancer antigen in studied serum samples were performed using a solid-phase sandwich CA 15-3 ELISA-BEST (FSSH-approved IVD medical device from Vector-BEST, Novosibirsk, Russia, Lot#T-8472) in accordance with the manufacturer's instruction. The CA 15-3-ELISA-BEST is designed to quantitatively measure the amount of cancer antigen CA 15-3 bound between a matched antibody pair in human serum. It uses two types of monoclonal antibodies specific to different CA 15-3 epitopes. Capture CA 15-3 specific monoclonal antibodies have been pre-coated in the wells of the supplied microplate. Samples, CA 15-3 standards, positive controls, and substrate solution for the secondary peroxidase HRP- labeled antibody are then added into the corresponding wells, allowed to react with the HRP-antibody-CA-15.3 complex to produce a measurable optical signal to be recorded with a colorimetric microplate reader. The sandwich is formed by the addition of the secondary antibody. In this case, the intensity of this signal is directly proportional to the concentration of CA 15-3 antigen present in the original specimen. This approach to sandwich ELISA allows the formation of the antibody-analyte sandwich complex in a single step. Time-to-result: 3.5 Hrs. Sensitivity of the assay: 0.5 U/ml, diagnostic range: 10 U/ml - 250 U/ml, detection method: colorimetric. For the optical density measurements was used a Tecan Spark (Tecan, Switzerland, EU). The assay had linearity ranges from 97 to 98% over the range of concentrations from 5 to 65 U/ml. The cut-off value (30 U/ml) was recommended by the manufacturer based on CA 15-3 levels in sera from healthy females (N=97) of 18 – 50 years old. All the test samples were done in duplicates using Standard calibrators CA 15-3 and Positive control sample as reference standards included in the set of reagents. The intra-assay coefficient of variation was $\leq 8.2\%$.

2.4 Statistical analysis

The distribution of CA-62 and CA 15-3 in sera from healthy and breast cancer patients was tested for normality using the D'Agostino-Pearson omnibus test. The Pearson correlation coefficient (r) was used

to determine the correlation between CA-62 and CA 15-3 serum levels. Since CA-62 values a 1000 times higher than the CA 15-3, the original values for both cancer markers in different subgroups were log-transformed (\log_{10}) before the analysis for obtaining the same equivalent scales, which allow getting a graphical correlation with $y = a \cdot \lg(x) + b$. For the evaluation of the difference between cancer and healthy control groups the Mann-Whitney U test was used. In order to evaluate the diagnostic characteristics of each cancer antigen we calculated the sensitivity and specificity, test accuracy, PPV, and NPV, and compared the cancer samples to normal control subjects using the receiver operating characteristic (ROC) analysis for the two markers. The level of significance was set at $p < 0.001$. Statistical analyses were performed using the MedCalc statistical software (version 19.7.4, MedCalc Software Ltd, Belgium, EU). The weighted kappa k-coefficients were used for evaluation of the diagnostic test results against a gold standard, which is, in our case, the results of the histopathological findings.

3 Results

Serum samples from 73 healthy control subjects (women) and 196 patients with histopathologically confirmed breast cancer were analyzed for CA-62 cancer marker and 177/196 (due to insignificant sample's volume) of breast cancer samples with 73 healthy control subjects for CA-15-3 serum levels, as described in the Materials and methods section. The values obtained are shown in Table 2. Significantly higher serum CA-62 levels were found in sera from breast cancer patients compared to healthy control women, and the glycoprotein concentration ranged from 1178 to 28598 U/ml (mean \pm SD=12312 \pm 5326) (Figure 2). The median CA-62 values were very high in all stages of breast cancer: ductal carcinoma *in situ* DCIS (12133 U/ml), Stage I (13045 U/ml), Stage II (9824 U/ml) and Stage III (17247 U/ml) as compared to healthy control subjects (2821 U/ml). Interestingly enough, the CA-62 detection level decreases with the tumor stage and demonstrates a very significant production of the marker for epithelial carcinomas from the onset of carcinogenesis, when cancer stem cells are poorly differentiated.

TABLE 2 Diagnostic methods comparison: CA 15-3 ELISA, CA-62 CLIA, mammography, MRI, UE, and MCED in relation to clinical and pathologic data of breast cancer patients.

Parameter	CA-62, U/ml Patient no (% > 5000)	CA15-3, U/ml Patient no (% > 30)	Mammograph [29, 36] vs DM, Patients % with positive results	MRI FAST ^[35] Patients % with positive results	UE Patients % with positive results	cfDNA based Multi-cancer early detection test MCED ^[30]
Stage DCIS (0)	14/14 (100%)	2/11 (18%)	36 vs 41%	80%	N/A	N/A
Stage I	95/98 (97%)	38/88 (43%)	34% vs 46.6%			16.8%
Stage II	61/73 (85%)	37/68 (54%)	80% - 90%			40.4%
Stage III	9/11 (83%)	5/10 (50%)	80% vs 92%			77%
Stage IV			92%			90.1%
Total	180/196(92%)	82/177 (46%)	63 vs 90%	64.8%	80%	51.5%
Sensitivity %	92%	42%	63-90% vs 97%	77 - 80%	81%	51.5%
Specificity %	93%	92%	60% vs 64.5%	64.2%	78.5%	99.5%
AUC	0.955	0.77	78% vs 89%	0.79	0.867	44.4%
PPV	97.8%	93%	80% vs 90.9%	N/A	75%	99.4%
NPV	81%	37%	75% vs 89.3%	N/A	85%	N/A
Test Accuracy%	92.2%	58%	83% vs 89%	75.5%	80.5%	N/A

*Published data based on accomplished screening programs.

*Using 0.7 breast cancer prevalence (USA and Western Europe) for population-based screening.

The CA 15-3 values in the breast cancer sera showed a broad range, from 0 to 330.4 U/ml with a median of 25.8 U/ml (mean \pm SD= 35 \pm 35.6) with its minimum (20.5 U/ml) at Stage I and its maximum (38.2 U/ml) at Stage III. A LOD value of 5 U/ml for the CA 15-3 ELISA immunoassay was used. The entire set of sera from healthy women had CA 15-3 antigen levels above the detection limit (5 U/ml) and 97% of them had detectable CA 15-3 levels in the range of 5 to 65 U/ml (the estimated mean \pm SD = 0.2 \pm 0.05). Less than half of women (46%) with confirmed breast cancer had elevated levels of CA 15-3 as compared to healthy control

subjects. D'Agostino-Pearson omnibus normality test for CA-62 levels in healthy women as well as in breast cancer patients showed Gaussian normal distributions, whereas cancer antigen CA 15-3 levels were not normally distributed for breast cancer patients as compared to healthy control subjects. Both CA-62 and CA 15-3 glycoprotein levels in healthy as well as in breast cancer patients did not have any significant correlation with the age of the individuals at 95% Confidence interval ($r = 0.129$, $p = 0.07$ for CA-62 and $r = 0.11$, $p = 0.16$). No correlation was found between the serum CA-62 and CA 15-3 levels in the healthy control group and in breast cancer

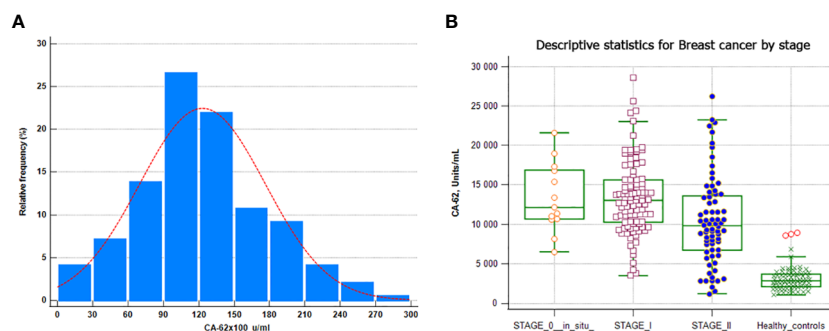


FIGURE 2

CA-62 normal distribution in breast cancer samples. (A) CA-62 levels in sera from healthy controls and from all breast cancer groups (B) Error bars denote maximum and minimum values.

patients ($r = 0.11$, $p < 0.0003$). Overall, a significant correlation was found between the CLIA-CA-62 assay and the histopathological findings ($r = 0.942$, $p < 0.0002$) using linear regression for the entire set of breast cancer and $r = 0.97$, $p < 0.0001$ for Stage I.

The overall performance of the competitive CLIA-CA-62 and CA 15-3 ELISA assays in sera from breast cancer patients with stages I-III and the ductal carcinomas *in situ* was evaluated by constructing ROC curves. The results of the ROC-curve analysis for Stage I and DCIS for both CA-62 and CA 15-3 are shown in Table 2 and Figure 3A and the ROC-curves for all stages are presented in Figure 3B. ROC curve analysis for the entire set of breast cancer samples using the CA-62 cancer marker demonstrated a very high AUC = 0.955 with $p < 0.001$. The CLIA-CA-62 assay for DCIS and Stage I breast cancer showed an AUC of 0.989 with $p < 0.0001$ using a cut-off value of 5000 U/ml recommended by the manufacturer of the assay; Sensitivity was 97% at 95% specificity (Figure 3A) with the median and an average equal to 12133 and 13062, correspondingly. By contrast, the CA 15-3 ELISA assay yielded an AUC = 0.779, $p < 0.001$ for the entire set of samples, and using a cut-off value of 30 U/ml, the Sensitivity was 46% at 93% Specificity, which corresponds well with previously published studies. The ROC-curve analysis of CA 15-3 ELISA assay for DCIS and Stage I breast cancer showed an AUC of 0.76, $p < 0.0001$ with a Sensitivity of 40% at the same 93% Specificity (Table 2; Figure 3). The comparison of the results between the two cancer markers revealed that the Sensitivity of the CLIA CA-62 chemiluminescent assay was approximately double that of CA 15-3 ELISA for early stages of breast cancer, and over three times as high in DCIS (27% vs. 100%). The accuracy of the test (the proportion of the correct test results in a total number of cases) among all examined patients using the CLIA-CA-62 assay is 97% for detecting Stage I breast cancer compared to 92% for the entire set, whereas for sandwich CA

15-3 ELISA it is only 40% (46% for the entire set). The positive predictive value (PPV) and negative predictive value (NPV) were used to describe the performance of the diagnostic test. Further analysis compared the CA 15-3 and serum CA-62 glycoprotein values with published results for other methods of cancer diagnostics such as mammography, ultrasound, and MRI. The Positive predictive value (PPV) for CLIA-CA-62 is 97.8% as compared to 93% for CA 15-3, 78% - 90% for mammography and 75% for MRI and ultrasound. At the same time, test accuracy is the highest at 92.2% for CLIA-CA-62 as compared to 58% for CA 15-3 ELISA, and 75-85% for mammography, ultrasound, and MRI.

4 Discussion

This study is unique in terms of the sensitivity demonstrated for the detection of the very early stages of breast cancer in asymptomatic women.

The aim of this study was not a differential diagnosis between breast cancer and breast benign disease using a biomarker CA-62 or an evaluation of the relationship between the level of CA-62 and different molecular subtypes. The main goal was to independently evaluate the diagnostic characteristics of the novel CLIA-CA-62 assay for early stages of breast cancer detection as compared to other CA 15-3 ELISA based on well-known cancer marker CA 15-3 and its prospective use thereof. However, the obtained results have demonstrated a potential in the future to carry out a prospective clinical study of the relationship between the CA-62 serum level and the tumor grade, as well as the molecular subtypes of breast cancer patients.

Measuring serum levels of cancer markers CA-62 and CA 15-3 in 269 samples has established that CA-62 antigen was increased in

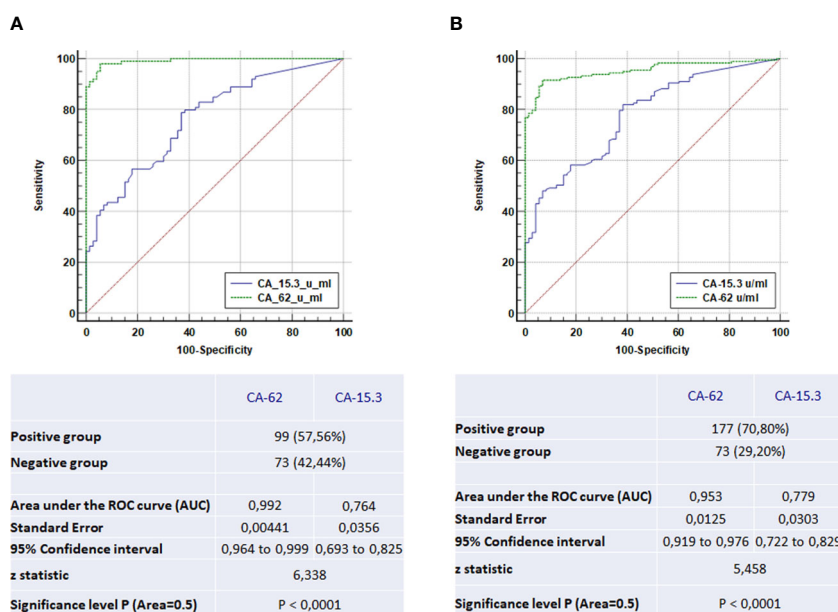


FIGURE 3

The ROC-curve comparison of the CLIA-CA-62 and CA 15-3 ELISA IVD assays for DCIS and Stage I breast cancer patients in comparison with healthy controls (A) and the ROC-curves comparison of the CLIA-CA-62 and CA 15-3 ELISA assays for all stages of breast cancer (B).

180/196 (92%) ($p < 0.0001$) of breast cancer patients in the DCIS, Stage I, Stage II, and Stage III, whereas elevated CA 15-3 values were found only in 46% (82/177). Both cancer markers in the healthy control subjects' group ($N=73$) had serum levels below the upper limit of the reference range (67/73 for CA-62 and 164/177 for CA 15-3). Cut-off values for CA 15-3 ELISA assay were > 30 U/ml, and > 5000 U/ml for CLIA-CA-62 upon which the sensitivity, specificity, AUC, and CI were calculated. Interestingly enough, patients with Stage I of breast cancer and non-invasive ductal carcinoma *in situ* DCIS have demonstrated from 97 to 100% detection using the novel CLIA-CA-62 assay, when curability is the highest, while the mucin-based CA 15-3 ELISA assay was not (27% detection only for DCIS). The same trend was observed for the entire set of breast cancer samples. Previous studies on cancer biomarkers for breast cancer (15) have demonstrated that their low sensitivity and specificity prevent from the use of serum markers such as the MUC-1 mucin glycoproteins (CA 15.3, BR 27.29) and carcinoembryonic antigen (CEA) for the diagnosis of early breast cancer. At the same time, serial measurement of these markers can result in the early detection of recurrent disease as well as reflect the efficacy of therapy.

The reason for such unusually high Sensitivity for early stages of BC detection found in this blind study could be that N-glycoprotein CA-62 is a mesenchymal light N-oligosaccharide, which is shed into the bloodstream far beyond the other cancer marker production. In general, epithelial tumors in a process of malignant transformation gradually lose their differentiation due to the destruction of the connections with the tumor microenvironment, which is controlling the degree of cell differentiation loss, up to the epithelial-mesenchymal transition with the formation of the tumor stem cells and re-expression of the embryonic antigens. During EMT epithelial cells lose their epithelial characteristics, and their polarity and gain some properties of the mesenchymal cells, such as spindle shape, anterior-posterior polarization, and strong migratory potential and mesenchymal markers. As a result, various epithelial cells possessing different morphological and phylogenetic classifications are transformed into the same pluripotent cancer stem cells. From the onset of carcinogenesis heavily branched N-glycoproteins are expressed in large quantities on the cell membrane of such "transformed" stem cells and act as a carcinoma-specific N-glycoprotein MEC/CA-62 detected by the CLIA-CA-62 chemiluminescent assay.

Authors (16, 17) revealed that a CA-62 marker associated with epithelial tumors can be significantly expressed and detected to varying degrees in the tissues of malignant tumors (such as breast, prostate, lung, uterus, stomach, kidney, colon, and ovaries), as well as in various human biological fluids (including blood and saliva). At the same time, healthy control subjects do not demonstrate increased expression of the CA-62 marker. However, some breast benign specimens ($<10\%$) have demonstrated a slight increase in CA-62 level that might indicate a transitional stage of the tumor becoming malignant, which was actually confirmed lately for some benign patients (18). Patients having a strong elevation in serum CA-62 level might have another type of carcinoma, which does not make it false positive for breast cancer detection, but rather a substantial reason for simultaneous detection of other existing cancer.

As compared to low-weight CA-62 N-glycoprotein, a majority of other cancer markers including CA 15-3 represents heavyweight O-mucins (up to 800 kDa), are getting produced when the cancer cells differentiation reaches maturity and are released into the blood after the tumor cells destruction. In this case, the level of released into the blood accumulated tumor-specific and tumor-associated markers is proportional to the tumor growth. That could be a reason for the low detection level of CA 15-3 in the serum of patients with non-invasive or micro-invasive breast lesions as compared to amounts seen in sera from patients with advanced cancers (9, 15). The accepted overall established sensitivity of the CA 15-3 assay for breast cancer detection is in the 20 to 50% range, which is in agreement with the findings reported herein ($Se = 27 - 56\%$) with lower values for DCIS (27%) and Stage I (42%), and higher values for Stage II (54%) and Stage III (50%) (23, 24).

According to the ASCO guidelines, CA 15-3 and CEA can be used only together with physical examination and imaging (23, 24), but mammography has a limited sensitivity of 63-90% due to many possible influences such as unclear lesions, poor aligning, dense parenchyma, calcifications, distortions, and misinterpretations (25). Hence, the probability that a patient with breast cancer will be detected by mammography alone is only 70-90%, and therefore the probability that it will be missed is 10 – 30%. Some multicenter trials established that up to 40% (30% at Stage I) of breast cancer cases are "missed cancers" due to detection and interpretation errors. At the same time, mammographic screening detects 50 per 100,000 (0.5%) invasive breast cancers in women while generating 2200 per 100,000 women (2.2%) false positive results from 90-99% of true negatives (26, 27).

Traditional bilateral mammography is being improved by using contrast-enhanced digital mammography (CEDM) which allows visualizing neovascularization associated with angiogenesis. Other greatly demanded and valuable screening methods are breast ultrasound and ultrasound elastography (UE), which shows superior advantages in differentiating benign and malignant breast tumors as compared to conventional ultrasound (28). On the other hand, breast MRI has some advantages over mammography since it does not use radiation, and is faster and exceptionally safe. MRI images reflect the tumor's molecular and genetic characteristics. Unlike mammography, which generates images based on the density of the tissue, MRI has higher sensitivity by creating a "blood flow map" which allows visualizing tumor neovascularity, associated with some metabolic modifications that correlate with the proliferation and metastatic potential of the tumor (29, 30).

The published results for various methods of cancer diagnostics such as mammography (MMG), magnetic resonance imaging (MRI), an ultrasound and ultrasound elastography (UE), and DNA-based Multi-cancer early detection test MCED (31) were compared with results obtained for 196 patients with breast cancer using immunological methods CLIA-CA-62 assay and CA 15-3 ELISA. A weighted kappa-test (k) was performed to evaluate the diagnostic consistency of the mammography, CA 15-3 ELISA and CLIA-CA-62 assays with the results of histopathology. The sensitivity and specificity values for the combination of the instrumental methods of diagnostics, such as UE & MRI were 95.8% and 92.8%, correspondingly, which are

comparable to the values obtained with the CLIA-CA-62 assay. Kappa coefficients are being interpreted as indices of the test quality in evidence-based medicine (32). In this case, the values of the weighed kappa-coefficient demonstrate a significant difference between the two diagnostic methods: 0.63 (UE&MRI) vs 0.80 (CLIA-CA-62). The comparison of the kappa coefficient (0.8) for the CLIA-CA-62 test with published elsewhere (28) kappa coefficients for UE, MRI, UE&MRI (0.512, 0.527 and 0.630, respectively, $p < 0.001$), for mammography (0.52), and CA 15-3 ELISA (0.22) with significance level $p < 0.0001$ indicate that the CLIA-CA-62 test classifies patients more reliably due to the lesser likelihood of a random coincidence of the test results with the histopathological findings.

Taken together, the results obtained in this independent double-blind study clearly demonstrate that the novel CLIA-CA-62 chemiluminescent assay has significant diagnostic advantages in detecting early stages of breast cancer as compared to other imaging diagnostic methods as well as to the other cancer markers including CA 15-3. The values of sensitivity, specificity, and accuracy of the CLIA-CA-62 IVD assay were 92%, 93%, and 92.2%, which is approximately 1.5 times higher as compared to various visual methods of diagnostics such as MRI, mammography, ultrasound etc. It is especially worth emphasizing the significant difference in detection of DCIS and Stage I & II. When using the CLIA-CA-62 assay, the detection rate was over 97%, whereas the other methods show a range of values from 27% for the CA15-3 assay (23, 24), to 55% for multi-cancer early detection test MCED (31) to 80% for FAST MRI (29).

The comparison of different breast cancer diagnostic methods allows concluding that only a combination of several methods is superior to the single use of either method for the detection of Stage I breast cancer. Data from Table 2 confirms that a combination of the CLIA-CA-62 and other methods of diagnostics including mammography could significantly improve the detection of non-invasive DCIS and Stage I breast cancer both having a high survival rate.

In the future, it seems appropriate to conduct a clinical approbation on a breast cancer screening of a group of women patients to develop a working algorithm for reliable detection of early-stage breast cancer using the CLIA-CA-62 immunoassay as a pre-screening tool before or in conjunction with breast mammography. It could be beneficial for current breast cancer screening algorithms and particularly for early-stage breast cancer detection. Another clinical study aiming the use of the CLIA-CA-62 immunoassay for differential diagnosis in women above 40 years old with BI-RADS 2, 3, and 4 mammograms, with moderate to high suspicion of breast cancer correlating mammographic and the CLIA-CA-62 data with pathologic findings might help with the interpretation of the pathologic findings and with the differentiation between benign lesions and malignant neoplasms.

Conclusions

1. The CLIA-CA-62 assay demonstrated 100% sensitivity at 93% specificity for DCIS and 97.8% for Stage I breast cancer

as compared to another known cancer marker for breast cancer, such as CA 15-3 (Se = 46%, Sp = 93%) and mammography (Se = 63-80% and 60% Specificity depending on the stage of cancer and parenchyma density).

2. Cancer marker CA-62 has a few unique qualities that distinguish it from other well-known cancer markers: it is present at a very high level in the blood of patients (>97%) with asymptomatic ductal carcinoma *in situ* and Stage I breast cancer and it doesn't increase along with the cancer progression and differentiation. Since the carcinoma-specific marker CA-62 appears on the surface of the transformed mesenchymal epithelial cancer cells in the course of carcinogenesis, it is gradually fading away, when tissue differentiation reaches maturity.
3. There is a significant level of agreement ($k=0.8$) between the CLIA-CA-62 assay results based on the marker for epithelial carcinomas CA-62 with histopathological findings for the entire set of breast cancer, including Stage I.
4. Despite the fact that the CA-62 is a marker for epithelial carcinomas and is not specific for breast cancer, the results obtained in this blind study suggest that the CLIA-CA-62 assay could be a useful tool to supplement existing mammography screening as well as other diagnostic imaging methods, which could improve the diagnostic sensitivity in DCIS and Stage I breast cancer detection thus improving clinical outcomes. Patients having a strong elevation in serum CA-62 level might have another type of carcinoma, which does not make it false positive for BC but rather benefits the simultaneous detection of some other existing cancer.
5. This evaluation of the CLIA-CA-62 chemiluminescent assay for early stages of breast cancer detection strongly suggests it can provide independent and complementary information for the doctors in decision-making and can be considered a useful tool for the primary detection of breast cancer in asymptomatic women. It would be beneficial to use serum CA-62 level in conjunction with the clinical information and other diagnostic procedures.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Local Ethics Committee of the Sechenov First Moscow State Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conception, MS. Interpretation or analysis of data, AB, AF, AR, and NB. Preparation of the manuscript, AB, AF, AR, and NB. Revision for important intellectual content, MS and AB. Supervision, MS. All authors contributed to the article and approved the submitted version.

Funding

This work was financed by the Ministry of Science and Higher Education of the Russian Federation within the framework of state support for the creation and development of World-Class Research Centers “Digital Biodesign and Personalized Healthcare” №075-15-2022-304.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
2. Lima SM, Kehm RD, Terry MB. Global breast cancer incidence and mortality trends by region, age-groups, and fertility patterns. *EClinicalMedicine* (2021) 38:100985. doi: 10.1016/j.eclinm.2021.100985
3. Kong Y-C, Bhoo-Pathy N, Subramaniam S, Bhoo-Pathy N, Taib NA, Jamaris S, et al. Advanced stage at presentation remains a major factor contributing to breast cancer survival disparity between public and private hospitals in a middle-income country. *Int J Environ Res Public Health* (2017) 14(4):427. doi: 10.3390/ijerph14040427
4. Rubinstein WS, Latimer JJ, Sumkin JH, Huerbin M, Grant SG, Vogel VG. Prospective screening study of 0.5 Tesla dedicated magnetic resonance imaging for the detection of breast cancer in young, high-risk women. *BMC Women's Health* (2006) 6:10. doi: 10.1186/1472-6874-6-10
5. Peintinger F. National breast screening programs across Europe. *Breast Care* (2019) 14:354–8. doi: 10.1159/000503715
6. Azadnajafabad S, Saeedi Moghaddam S, Keykhaei M, Shobeiri P, Rezaei N, Ghasemi E, et al. Expansion of the quality of care index on breast cancer and its risk factors using the global burden of disease study. *Cancer Med* (2022) 00:1–15. doi: 10.1002/cam4.4951
7. da Costa Vieira RA, Biller G, Uemura G, Ruiz CA, Curado MP. Breast cancer screening in developing countries. *Clinics (Sao Paulo)* (2017) 72(4):244–53. doi: 10.6061/clinics/2017(04)09
8. *AJCC cancer staging manual*. In Amin MB ed. American College of Surgeons (2017).
9. Jainish P, Pritesh P. Biosensors and biomarkers: promising tools for cancer diagnosis. *Int J Biosen Bioelectron* (2017) 3(4):313–6. doi: 10.15406/ijbsbe.2017.03.00072
10. Duffy MJ, Walsh S, McDermott EW, Crown J. Biomarkers in breast cancer: where are we and where are we going? *J Adv Clin Chem* (2015) 71:1–23. doi: 10.1016/bs.acc.2015.05.001
11. Voronova V, Glybochko P, Svistunov A, Fomin V, Kopylov P, Tzarkov P, et al. Diagnostic value of combinatorial markers in colorectal carcinoma. *Front Oncol* (2020) 10:832. doi: 10.3389/fonc.2020.00832
12. Voronova V, Peskov K, Glybochko P, Svistunov A, Fomin V, Kopylov P, et al. Evaluation and diagnostic potential of plasma biomarkers in bladder cancer. *Ann Oncol* (2019) 30(5):V45–7. doi: 10.1093/annonc/mdz239
13. Glybochko PV, Svistunov A, Fomin V, Kopylov P, Sekacheva M, Vasiliev I, et al. Method for screening probability of breast cancer presence. Patent EA202090713A2 (2020). Available at: https://ru.espacenet.com/publicationDetails/originalDocument?FT=D&date=20210129&DB=EPODOC&locale=ru_RU&CC=EA&NR=202090713A3&KC=A3&ND=4
14. Bhushan A, Gonsalves A, Menon JU. Current state of breast cancer diagnosis, treatment, and theranostics. *Pharmaceutics* (2021) 13:723–30. doi: 10.3390/pharmaceutics13050723
15. Park BW, Oh JW, Kim JH, Park SH, Kim KS, Kim JH, Lee KS. Preoperative CA 15-3 and CEA serum levels as predictor for breast cancer outcomes. *Ann Oncol Eur Soc Med Oncol (ESMO)* (2008) 19(4):675–81. doi: 10.1093/annonc/mdm538

Conflict of interest

Author NB was employed by B.V. Petrovsky Russian Scientific Center of Surgery.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

16. Cherkasova JR, Tsurkan SA, Kondratiev VB, Moro-Vidal R. *Cancer antigen for early cancer detection*//Patent RU202114411A, WO2021215955A1 (2021). Available at: <https://patents.google.com/patent/WO2021215955A1/en?q=WO2021215955A1>.
17. Tcherkassova J, Prostyakova A, Tsurkan S, Ragoulin V, Boroda A, Sekacheva M, et al. Diagnostic efficacy of the new prospective biomarker's combination CA 15-3 and CA-62 for early-stage breast cancer detection: results of the blind prospective-retrospective clinical study. *Cancer Biomarkers* (2022) 35(1):57–69. doi: 10.3233/CBM-210533
18. Cherkasova Zh. R, Tsurkan SA, Kondratiev VB. *Set of reagents for detecting the marker for epithelial carcinomas*//Patent RU2735918C2, US2022057401A1 (2022). Available at: <https://worldwide.espacenet.com/patent/search/family/062981387/publication/US2022057401A1?q=US2022057401A1>.
19. Khakimova GG, Cherkasova Z, Tsurkan SA, Fedchikov GA, Suganov NV, Gorbunova VA. A pilot clinical trial to monitor response to chemotherapy using the CA-62 marker of epithelial carcinomas. *Siberian J Oncol* (2019) 18(5):18–28. doi: 10.21294/1814-4861-2019-18-5-18-28
20. Sawaki M, Shien T, Iwata H. TNM classification of malignant tumors (Breast cancer study group). *Jpn J Clin Oncol* (2019) 49(3):228–31. doi: 10.1093/jjco/hyy182
21. Leclercq I, Batejat C, Burguière AM, Manuguerra JC. Heat inactivation of the middle East respiratory syndrome coronavirus. *Influenza Other Respir Viruses* (2014) 8(5):585–6. doi: 10.1111/irv.12261
22. Tuck MK, Chan DW, Chia D, Godwin AK, Grizzle WE, Krueger KE, et al. Standard operating procedures for serum and plasma collection: early detection research network consensus statement standard operating procedure integration working group. *J Proteome Res* (2009) 8(1):113–7. doi: 10.1021/pr800545q
23. Duffy MJ. Serum tumor markers in breast cancer: are they of clinical value? *Clin Chem* (2006) 52(3):345–51. doi: 10.1373/clinchem.2005.059832
24. Duffy MJ, Shering S, Sherry F, McDermott E, O'Higgins N. CA 15-3: a prognostic marker in breast cancer. *Int J Biol Markers* (2000) 15(4):330–3. doi: 10.1177/17246080001500410
25. Majid AS, de Paredes ES, Doherty RD, Sharma NR, Salvador X. Missed breast carcinoma: pitfalls and pearls. *Radiographics* (2003) 23(4):881–95. doi: 10.1148/rg.234025083
26. Giampietro RR, Cabral MVG, Lima SAM, Weber SAT, Dos Santos Nunes-Nogueira V. Accuracy and effectiveness of mammography versus mammography and tomosynthesis for population-based breast cancer screening: a systematic review and meta-analysis. *Sci Rep* (2020) 10(1):7991. doi: 10.1038/s41598-020-64802-x
27. Hofvind S, Geller BM, Skelly J, Vacek PM. Sensitivity and specificity of mammographic screening as practiced in Vermont and Norway. *Br J Radiol* (2012) 85(1020):e1226–32. doi: 10.1259/bjr/15168178
28. Cheng R, Li J, Ji L, Liu H, Zhu L. Comparison of the diagnostic efficacy between ultrasound elastography and magnetic resonance imaging for breast masses. *Exp Ther Med* (2018) 15(3):2519–24. doi: 10.3892/etm.2017.5674
29. Morris EA. Rethinking breast cancer screening: ultra FAST breast magnetic resonance imaging. *J Clin Oncol* (2014) 32(22):2281–3. doi: 10.1200/JCO.2014.56.1514
30. Zeeshan M, Salam B, Khalid QSB, Alam S, Sayani R. Diagnostic accuracy of digital mammography in the detection of breast cancer. *Cureus* (2018) 10(4):e2448. doi: 10.7759/cureus.2448

31. Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. *Ann Oncol* (2021) 32(9):1167–77. doi: 10.1016/j.annonc.2021.05.806
32. Gilchrist JM. Weighted 2 x 2 kappa coefficients: recommended indices of diagnostic accuracy for evidence-based practice. *J Clin Epidemiol* (2009) 62(10):1045–53. doi: 10.1016/j.jclinepi.2008.11.012



OPEN ACCESS

EDITED BY

Laura Curiel,
University of Calgary, Canada

REVIEWED BY

Subathra Adithan,
Jawaharlal Institute of Postgraduate
Medical Education and Research
(JIPMER), India
Hanxiong Guan,
Huazhong University of Science and
Technology, China

*CORRESPONDENCE

Meiqin Wang
✉ meiqin-wang@163.com
Zhengyang Zhou
✉ glzyzhou@163.com
Wenrong Shen
✉ jszlyyct@126.com

[†]These authors have contributed equally to
this work

RECEIVED 18 January 2023

ACCEPTED 25 April 2023

PUBLISHED 05 May 2023

CITATION

Xie X, Liu K, Luo K, Xu Y, Zhang L, Wang M,
Shen W and Zhou Z (2023) Value of dual-
layer spectral detector computed
tomography in the diagnosis of benign/
malignant solid solitary pulmonary nodules
and establishment of a prediction model.
Front. Oncol. 13:1147479.
doi: 10.3389/fonc.2023.1147479

COPYRIGHT

© 2023 Xie, Liu, Luo, Xu, Zhang, Wang, Shen
and Zhou. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Value of dual-layer spectral detector computed tomography in the diagnosis of benign/malignant solid solitary pulmonary nodules and establishment of a prediction model

Xiaodong Xie^{1,2†}, Kaifang Liu^{1†}, Kai Luo¹, Youtao Xu³, Lei Zhang¹,
Meiqin Wang^{1*}, Wenrong Shen^{1*} and Zhengyang Zhou^{2*}

¹Department of Radiology, Nanjing Medical University Affiliated Cancer Hospital, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, Nanjing, China, ²Department of Radiology, Nanjing Drum Tower Hospital, Clinical College of Nanjing Medical University, Nanjing, China, ³Department of Thoracic Surgery, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, Nanjing Medical University Affiliated Cancer Hospital, Nanjing, China

Objective: This study aimed to investigate the role of spectral detector computed tomography (SDCT) quantitative parameters and their derived quantitative parameters combined with lesion morphological information in the differential diagnosis of solid SPNs.

Methods: This retrospective study included basic clinical data and SDCT images of 132 patients with pathologically confirmed SPNs (102 and 30 patients in the malignant and benign groups, respectively). The morphological signs of SPNs were evaluated and the region of interest (ROI) was delineated from the lesion to extract and calculate the relevant SDCT quantitative parameters, and standardise the process. Differences in qualitative and quantitative parameters between the groups were statistically analysed. A receiver operating characteristic (ROC) curve was constructed to evaluate the efficacy of the corresponding parameters in the diagnosis of benign and malignant SPNs. Statistically significant clinical data, CT signs and SDCT quantitative parameters were analysed using multivariate logistic regression to determine the independent risk factors for predicting benign and malignant SPNs, and the best multi-parameter regression model was established. Inter-observer repeatability was assessed using the intraclass correlation coefficient (ICC) and Bland–Altman plots.

Results: Malignant SPNs differed from benign SPNs in terms of size, lesion morphology, short spicule sign, and vascular enrichment sign ($P < 0.05$). The SDCT quantitative parameters and their derived quantitative parameters of malignant SPNs (SAR_{40keV}, SAR_{70keV}, Δ_{40keV} , Δ_{70keV} , CER_{40keV}, CER_{70keV}, NEF_{40keV}, NEF_{70keV}, λ , NIC, NZ_{eff}) were significantly higher than those of benign SPNs ($P < 0.05$). In the subgroup analysis, most parameters could distinguish

between benign and adenocarcinoma groups (SAR_{40keV} , SAR_{70keV} , Δ_{40keV} , Δ_{70keV} , CER_{40keV} , CER_{70keV} , NEF_{40keV} , NEF_{70keV} , λ , NIC, and NZ_{eff}), and between benign and squamous cell carcinoma groups (SAR_{40keV} , SAR_{70keV} , Δ_{40keV} , Δ_{70keV} , NEF_{40keV} , NEF_{70keV} , λ , and NIC). However, there were no significant differences between the parameters in the adenocarcinoma and squamous cell carcinoma groups. ROC curve analysis indicated that NIC, NEF_{70keV} , and NEF_{40keV} had higher diagnostic efficacy for differentiating benign and malignant SPNs (area under the curve [AUC]: 0.869, 0.854, and 0.853, respectively), and NIC was the highest. Multivariate logistic regression analysis showed that size (OR=1.138, 95% CI 1.022–1.267, $P=0.019$), Δ_{70keV} (OR=1.060, 95% CI 1.002–1.122, $P=0.043$), and NIC (OR=7.758, 95% CI 1.966–30.612, $P=0.003$) were independent risk factors for the prediction of benign and malignant SPNs. ROC curve analysis showed that the AUC of size, Δ_{70keV} , NIC, and a combination of the three for differential diagnosis of benign and malignant SPNs were 0.636, 0.846, 0.869, and 0.903, respectively. The AUC for the combined parameters was the largest, and the sensitivity, specificity, and accuracy were 88.2%, 83.3% and 86.4%, respectively. The SDCT quantitative parameters and their derived quantitative parameters in this study exhibited satisfactory inter-observer repeatability (ICC: 0.811–0.997).

Conclusion: SDCT quantitative parameters and their derivatives can be helpful in the differential diagnosis of benign and malignant solid SPNs. The quantitative parameter, NIC, is superior to the other relevant quantitative parameters and when NIC is combined with lesion size and Δ_{70keV} value for comprehensive diagnosis, the efficacy could be further improved.

KEYWORDS

dual-layer spectral detector computed tomography, solitary pulmonary nodule, quantitative parameters, differential diagnosis, prediction model

1 Introduction

In 2020, there were approximately 2.2 million new cases of lung cancer worldwide, accounting for 11.4% of all malignant tumours, and approximately 1.8 million deaths, accounting for 18.0% of all deaths due to malignant tumours (1, 2). The 5-year survival rate of patients with stage I lung cancer is 60%–80%, while that of stage IV patients is only 5.3%. Therefore, early detection, diagnosis, and treatment are important for patients with lung cancer (3). The common manifestations in the early stages of lung cancer, solitary pulmonary nodules (SPNs) are round or round-like pulmonary parenchymal lesions with relatively clear boundaries and a diameter of ≤ 30 mm (4). The detection of SPNs is increasing owing to the popularisation of CT examinations (5). However, due to the atypical

clinical symptoms and incidental nature in some patients, SPNs can manifest as primary lung cancer, pulmonary metastatic nodules, infection foci, or other benign lesions. In addition, the malignancy rate of SPNs is only 30%–40% (6). Therefore, quick and accurate diagnosis of malignant SPNs and the provision of standardised diagnosis and treatment guidance are challenging.

Detection of SPNs by low-dose spiral computed tomography (LDCT) screening and early treatment interventions decreases the mortality of lung cancer patients by 20% (7); however, LDCT plays little role in the qualitative diagnosis of SPNs (8). Conventional CT plays an important role in evaluating the radiological signs and enhancement characteristics of SPNs, but some benign and malignant SPNs have highly similar lesion morphologies and enhancement patterns, and thus pose a great challenge in differentiating benign and malignant SPNs (9). ^{18}F -fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F -FDG PET/CT) can reflect the internal metabolism of SPNs based on conventional CT, but false positive results due to inflammation, infection and other factors are difficult to avoid (10). In addition, the high radiation dose and the expensiveness of ^{18}F -FDG PET/CT limits its application and promotion in China. Needle biopsy has high accuracy in differentiating benign from malignant SPNs. However, as an

Abbreviations: SDCT, Spectral detector computed tomography; SPNs, Solitary pulmonary nodules; ROI, Region of interest; ROC, Receiver operating characteristic; IC, Iodine concentration; ICC, Intraclass correlation coefficient; CT, Computed tomography; AUC, Area under the curve; ^{18}F -FDG PET/CT, ^{18}F -fluorodeoxyglucose positron emission tomography/computed tomography; ED, Electron density; MPR, Multi-planar reconstruction techniques; VNC, Virtual non-contrast; NIC, Normalised iodine concentration; NZ_{eff} , Normalised effective atomic number; NED, Normalised electron density.

invasive examination, patients not only bear the risk of complications (pneumothorax, bleeding, infection, needle dissemination, and implantation) but also economic and mental stress. Therefore, it is very important and urgent to explore a non-invasive and accurate method for differentiating benign from malignant SPNs.

Dual-layer spectral detector computed tomography (SDCT), a milestone in the development of CT, uses materials to produce different energy information at various X-ray energies to convert simple CT images to functional imaging. In contrast to conventional CT, SDCT can also obtain quantitative parameters such as iodine concentration (IC), the slope of the spectral attenuation curve (λ), Z-effective (Z_{eff}), and electron density (ED), in addition to CT values. Currently, SDCT is gradually being applied and promoted in many clinical fields (11, 12). In clinical oncology, SDCT is recognised for its ability to differentiate benign and malignant primary lesions, malignant tumour subtypes, primary lesions from metastases, and lymph node metastasis (LNM) predictions (13–15). Nonetheless, studies focusing on the differential diagnoses of benign and malignant pulmonary nodules are limited. Published literature focused on the quantitative parameters of SDCT in the differential diagnosis of benign and malignant SPNs (16–18). However, the clinical data of the patients and the morphological characteristics of the tumour were not included in the comprehensive analysis. Hence, this study aimed to 1) explore the importance of SDCT quantitative parameters and their derivatives, combined with morphological information of the lesion and clinical data of patients in the differential diagnosis of benign and malignant solid SPNs and 2) establish a prediction model for malignant SPNs to assist in the development of its best treatment strategy.

2 Materials and methods

2.1 Study population

The study was approved by the institutional ethics committee of Jiangsu Cancer Hospital (NO.ZN202212, Nanjing, Jiangsu, China) and the requirement for individual consent for this retrospective analysis was waived.

Data from 207 patients with newly diagnosed SPNs who underwent spectral CT examination in our hospital between September 2021 and August 2022 were retrospectively collected. The inclusion criteria for the participants were as follows: (i) SDCT chest scan performed in our hospital; (ii) patients diagnosed with solid SPNs with a lesion diameter of > 8 mm and ≤ 30 mm; and (iii) definite pathological results confirmed by surgery or puncture. The exclusion criteria were: (i) plain chest scan without enhanced scan ($n = 52$); (ii) anti-tumour therapy before SDCT examination ($n = 11$); (iii) patients with more than one solid solitary nodules ($n = 9$); and (iv) poor image quality ($n = 3$). Finally, 132 patients were enrolled in this study, including 102 with malignant SPNs and 30 with benign SPNs. A flow chart of patient inclusion and exclusion is shown in Figure 1.

Based on tumour pathology, the study participants were divided into two groups: malignant ($n = 102$) and benign ($n = 30$). The

histopathological subtypes of malignant SPNs of the patients included are adenocarcinoma ($n = 76$); squamous cell carcinoma ($n = 21$); small cell lung cancer ($n = 4$); adenosquamous carcinoma ($n = 1$). Among 30 patients with benign SPNs, inflammation ($n = 9$), pulmonary hamartoma ($n = 8$), pulmonary tuberculosis ($n = 8$), inflammatory pseudotumour ($n = 2$), alveolar cytoma ($n = 1$), congenital cystic adenomatoid malformation ($n = 1$), and collagen nodules with lymphoproliferation ($n = 1$) were observed.

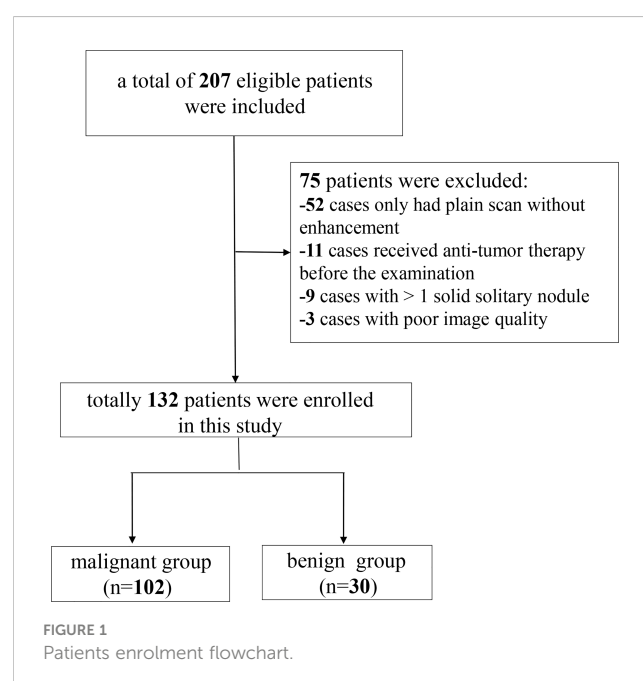
2.2 SDCT image acquisition

All patients underwent the same routine SDCT protocol (IQon; Philips Healthcare, Best, Netherlands). The patients were placed in the supine position, and the scanning range was from the thoracic entrance to the level of the costophrenic angle. The basic scanning parameters were: tube voltage, 120 kVp; tube current modulation; 3D modulation; collimator width, 64×0.625 mm; matrix, 512×512 ; and scanning field of view, 372 mm; pitch, 0.90; rotation time, 0.50 s. The slice thickness was 5 mm for scanning and 1 mm for reconstruction. Contrast medium (ioversol, 3.0 ml/kg; iodine, 350 mg/ml; HengRui Medicine, Jiangsu, China) was injected intravenously in the anterior elbow at a flow rate of 2.5–3.0 ml/s, followed by 20 ml normal saline at the same flow rate. A contrast-enhanced chest scan was performed 50 s later after the injection.

2.3 Radiological analysis of SPNs on SDCT

2.3.1 Analysis of the morphological characteristics

The images of all SPNs cases were interpreted and analysed by two radiologists (with 4 years and 8 years of work experience) blinded to the pathological findings. For analysing the morphological features of



SPNs, different window widths, window levels, thin-slice scans, and multi-planar reconstruction techniques were used to observe the lesion location, shape, lobulation, spiculation, pleural indentation, vascular convergence sign, air bronchogram, vacuole sign, calcification, and other imaging signs. When perceptions were inconsistent, a consensus was reached through consultation.

2.3.2 Acquisition and analysis of SDCT quantitative parameters

All images were imported into a Philips workstation (IntelliSpace Portal, Philips Healthcare), and the analysis and processing of the images were performed using the software in the workstation (Spectral CT Viewer, Philips Healthcare). Image analysis was performed by a radiologist (with 4 years of radiology experience) and supervised by a senior radiologist (with 8 years of radiology experience). All SPNs image data were acquired within the mediastinal window. Regions of interest (ROI) were manually delineated at three consecutive levels, including the largest level of the SPNs and its adjacent upper and lower levels. Areas such as calcification, vessels, cavities, atelectasis, and necrosis, which may affect measurements, were avoided. An average of three measurements for each case was taken as the final data point for the analysis. The ROI of the aorta at similar levels was obtained using the same method and used for the normalisation of SDCT quantitative parameters. Subsequently, a series of parameters was obtained.

The parameters obtained in our study were: CT values of SPNs at virtual non-contrast (VNC) and enhancement (40keV and 70keV) were recorded as $CT_{SPN-VNC}$, $CT_{SPN-40keV}$ and $CT_{SPN-70keV}$, respectively. The CT values of the aorta at the VNC ($CT_{aorta-VNC}$) and enhancement at 70keV (CT_{aorta}) were recorded and used as a reference. The following formulas were used to calculate the ratio of lesion to aortic virtual plain scan (SAR_{VNC}), the ratio of the lesion to aortic contrast enhancement (SAR_{40keV} , SAR_{70keV}), the difference in CT value between lesion enhanced and virtual plain scan (Δ_{40keV} , Δ_{70keV}), and the contrast enhancement ratio (CER_{40keV} , CER_{70keV}). Standardised enhancement value scores (NEF_{40keV} , NEF_{70keV}) and spectral curve slopes (λ) (19).

$$SAR_{VNC} = CT_{SPN-VNC} / CT_{aorta-VNC}$$

$$SAR_{40keV/70keV} = CT_{SPN-40keV/70keV} / CT_{aorta}$$

$$\Delta_{40keV/70keV} = CT_{SPN-40keV/70keV} - CT_{SPN-VNC}$$

$$CER_{40keV/70keV} = \Delta_{40keV/70keV} / CT_{SPN-VNC}$$

$$NEF_{40keV/70keV} = \Delta_{40keV/70keV} / (CT_{aorta} - CT_{aorta-VNC})$$

$$\lambda = (CT_{SPN-40keV} - CT_{SPN-70keV}) / (70 - 40)$$

To account for the haemodynamic variation between patients, the IC (mg/mL), Z_{eff} and ED values were normalised to the aorta. The normalised iodine concentration (NIC), normalised electron density (NED), and normalised effective atomic number (NZ_{eff}) were calculated according to the following formula:

$$NIC = IC_{SPN} / IC_{aorta}$$

$$NED = NED_{SPN} / NED_{aorta}$$

$$NZ_{eff} = Z_{eff-SPN} / Z_{eff-aorta}$$

To assess inter-observer reproducibility and variability, 50% of the study participants (66/132) were randomly selected, and the previous measurement procedure was repeated by another radiologist. Interobserver agreement was assessed using Bland–Altman plots and intraclass correlation coefficients (ICC).

2.4 Statistical analyses

The data were statistically analysed using SPSS 22 (SPSS, Inc., Chicago, IL, USA) and MedCalc15 (MedCalc Software, Mariakerke, Belgium). Continuous variables were presented as mean \pm standard deviation. The basic clinical data and SDCT morphological characteristics were analysed using the chi-square test or Fisher's exact test. The SPNs SDCT quantitative parameters and derived parameters were tested for normality using the Kolmogorov–Smirnov method, followed by the Levene test for homogeneity of variance. An independent sample t-test or Mann–Whitney U test was used to compare multiparameter differences between benign and malignant SPNs. One-way Analysis of Variance (ANOVA) was used to compare the pairwise differences between the benign and adenocarcinoma groups, benign and squamous cell carcinoma groups, and adenocarcinoma and squamous cell carcinoma groups. Receiver operating characteristic (ROC) curve analysis was used to determine the area under the curve (AUC), accuracy, sensitivity, and specificity of the different parameters in predicting benign and malignant SPNs. The clinical data of SPNs, SDCT morphological features, SDCT quantitative parameters, and their derived parameters with statistical differences were analysed by multivariate logistic regression to determine the independent risk factors for predicting benign and malignant SPNs, and the best multi-parameter regression prediction model was established. The inter-observer agreement of spectral CT parameters was evaluated using ICC (0.000–0.200: poor; 0.201–0.400: general; 0.401–0.600: medium; 0.601–0.800: good; 0.801–1.000: excellent) and the Bland–Altman plot evaluation.

3 Results

3.1 Basic clinical data of the study participants

Of the 132 pathologically confirmed solid SPNs patients included, 102 showed malignancy (42 males; 59 females) with an average age of 64.147 ± 9.491 years, and 61 of them were smokers. The 30 patients with benign tumours (14 males; 16 females) had an average age of 60.233 ± 10.047 years, and 12 of them were smokers. There were no significant differences in age ($P = 0.052$), sex ($P = 0.279$), or smoking history ($P = 0.985$) between the benign and malignant groups [Table 1](#).

TABLE 1 Basic clinical data of the patients.

	Benign group (n=30)	Malignant group (n=102)	P-value
Age (years)	60.233 ± 10.047	64.147 ± 9.491	0.052
Sex			0.279
male	14	43	
female	16	59	
Smoking			0.985
yes	12	61	
no	18	41	

3.2 Morphological features of benign and malignant SPNs

Table 2 shows the differences in the morphological features of benign and malignant SPNs identified using SDCT. The radiological findings of the two groups were analysed based on the following 10 aspects: location, size, lesion shape, lobulation, short spiculation, pleural indentation, vascular convergence sign, air bronchogram, vacuole sign, and calcification. There were significant differences in lesion size ($P = 0.04$), shape ($P = 0.014$), short spiculation sign ($P = 0.034$), and vascular convergence sign ($P = 0.015$) between the benign and malignant groups. The difference in the location, lobulation sign, pleural indentation sign, air bronchogram sign, vacuole sign, and calcification were not significant between the two groups ($P > 0.05$).

3.3 Comparison of SDCT quantitative parameters between the benign and malignant groups

The SDCT quantitative parameters were obtained by ROI delineation and then calculated. 13 quantitative parameter indices were included in this study for statistical analysis, and the specific quantitative parameter analysis findings are shown in Table 3. Except for SAR_{VNC} and NED ($P = 0.114$ and 0.208 , respectively), all other parameters (SAR_{40keV}, SAR_{70keV}, Δ_{40keV} , Δ_{70keV} , CER_{40keV}, CER_{70keV}, NEF_{40keV}, NEF_{70keV}, λ , NIC, and NZ_{eff}) were significantly higher in the malignant group than in the benign group ($P < 0.05$). A schematic diagram of ROI delineation and quantitative parameter acquisition for typical cases of benign and malignant SPNs are shown in Figures 2 and 3, respectively.

3.4 Differences in SDCT quantitative parameters among benign, adenocarcinoma, and squamous cell carcinoma groups

The SDCT quantitative parameters among the benign, adenocarcinoma, and squamous cell carcinoma groups are shown

in Table 4. Except for SAR_{VNC} and NED ($P = 0.788$ and 0.572 , respectively), the other parameters (SAR_{40keV}, SAR_{70keV}, Δ_{40keV} , Δ_{70keV} , CER_{40keV}, CER_{70keV}, NEF_{40keV}, NEF_{70keV}, λ , NIC, and NZ_{eff}) were significantly higher in the adenocarcinoma group than in the benign group ($P < 0.05$). Similarly, except for SAR_{VNC}, CER_{40keV}, CER_{70keV}, NED, and NZ_{eff} ($P = 0.209$, 0.434 , 0.230 , 0.404 , and 0.237), the other eight parameters (SAR_{40keV}, SAR_{70keV}, Δ_{40keV} , Δ_{70keV} , NEF_{40keV}, NEF_{70keV}, λ , and NIC) were significantly higher in the squamous cell carcinoma group than in the benign group ($P < 0.05$). There were no significant differences in the SDCT quantitative parameters between the adenocarcinoma and squamous cell carcinoma groups ($P > 0.05$).

3.5 Diagnostic performance of SDCT quantitative parameters

The diagnostic performance of SDCT quantitative parameters for different groups is shown in Tables 5–7 and Figure 4. Among the multiple quantitative parameters, NIC, NEF_{70keV} and NEF_{40keV} had the highest diagnostic efficiency in differentiating benign from malignant SPNs (AUC = 0.869, 0.854, and 0.853, respectively), and the AUC of NIC was the highest. When the cut-off value of NIC was 0.165, the sensitivity was 87.3%, the specificity was 76.7%, the positive prediction rate was 92.7%, the negative prediction rate was 63.9%, and the accuracy was 84.2%. NIC, Δ_{70keV} and NEF_{70keV} had the highest diagnostic efficiency in distinguishing benign SPNs from adenocarcinoma (AUC = 0.876, 0.857, and 0.854, respectively) and the NIC exhibited the highest detection efficiency. When the cut-off value of the NIC was 0.158, the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were 92.1%, 73.3%, 89.7%, 78.6%, and 86.0%, respectively. The three parameters with the highest diagnostic efficiency to distinguish between benign SPNs and squamous cell carcinoma were SAR_{40keV}, NEF_{40keV} and NEF_{70keV} (AUC = 0.849, 0.849, and 0.837, respectively). The diagnostic efficiency of the NIC was 0.832, which was second only to the above three parameters. In general, NIC is the best quantitative index for differentiating benign from malignant SPNs and for the comparison between the different subgroups.

TABLE 2 Radiological morphological signs of benign and malignant SPNs on SDCT images.

	Benign group (n=30)	Malignant group (n=102)	P-value
Location (lung)			0.496
upper lobe of left	9	24	
lower lobe of left	6	14	
upper lobe of right	6	38	
middle lobe of right	3	7	
lower lobe of right	6	19	
size	15.772 ± 5.661 mm	18.225 ± 5.689 mm	0.040*
Lesion shape			0.014*
round	10	14	
irregular	20	88	
Lobulation			0.123
yes	17	73	
no	13	29	
Short spiculation			0.034*
yes	9	53	
no	21	49	
Pleural indentation			0.095
yes	6	37	
no	24	65	
Vascular convergence sign			0.015*
yes	8	53	
no	22	49	
Air bronchogram sign			0.053
yes	5	36	
no	25	66	
Vacuole sign			0.296
yes	1	12	
no	29	90	
Calcification			0.694
yes	3	7	
no	27	95	

*p value<0.05.

3.6 Morphological features in combination with SDCT quantitative parameters for the differential diagnosis of benign and malignant SPNs

Binary logistic regression analysis performed with the SPNs size, shape, short spiculation, vascular convergence, and 11 quantitative parameters (SAR_{40keV} , SAR_{70keV} , Δ_{40keV} , Δ_{70keV} , CER_{40keV} , CER_{70keV} , NEF_{40keV} , NEF_{70keV} , λ , NIC, and NZ_{eff}) that were different between

benign and malignant groups as independent variables, and benign or malignant SPNs as the dependent variables showed that lesion size (OR = 1.138, 95% CI 1.022-1.267, $P=0.019$), Δ_{70keV} (OR=1.060, 95% CI 1.002-1.122, $P=0.043$), and NIC (OR=7.758, 95% CI 1.966-30.612, $P=0.003$) were independent risk factors for the prediction of benign and malignant SPNs. ROC curve analysis showed that the AUC of size, Δ_{70keV} , NIC, and a combination of the three for differential diagnosis of benign and malignant SPNs were 0.636, 0.846, 0.869, and 0.903, respectively. The AUC of the combination of the parameters (size

TABLE 3 Comparison of SDCT quantitative parameters between benign and malignant SPNs.

Parameters	Benign group (n=30)	Malignant group (n=102)	P-value
SAR _{VNC}	0.707 ± 0.357	0.785 ± 0.229	0.114
SAR _{40keV}	0.512 ± 0.198	0.841 ± 0.267	<0.001*
SAR _{70keV}	0.240 ± 0.098	0.351 ± 0.102	<0.001*
Δ _{40keV}	83.884 ± 43.997	146.955 ± 46.532	<0.001*
Δ _{70keV}	21.944 ± 13.584	42.059 ± 14.872	<0.001*
CER _{40keV}	3.262 ± 2.480	5.241 ± 3.577	<0.001*
CER _{70keV}	0.830 ± 0.699	1.453 ± 1.035	<0.001*
NEF _{40keV}	0.462 ± 0.212	0.872 ± 0.364	<0.001*
NEF _{70keV}	0.120 ± 0.067	0.247 ± 0.110	<0.001*
λ	1.990 ± 1.091	3.489 ± 1.141	<0.001*
NIC	0.134 ± 0.064	0.260 ± 0.101	<0.001*
NED	0.961 ± 0.089	0.981 ± 0.011	0.208
NZ _{eff}	0.764 ± 0.040	0.804 ± 0.080	<0.001*

*p value<0.05.

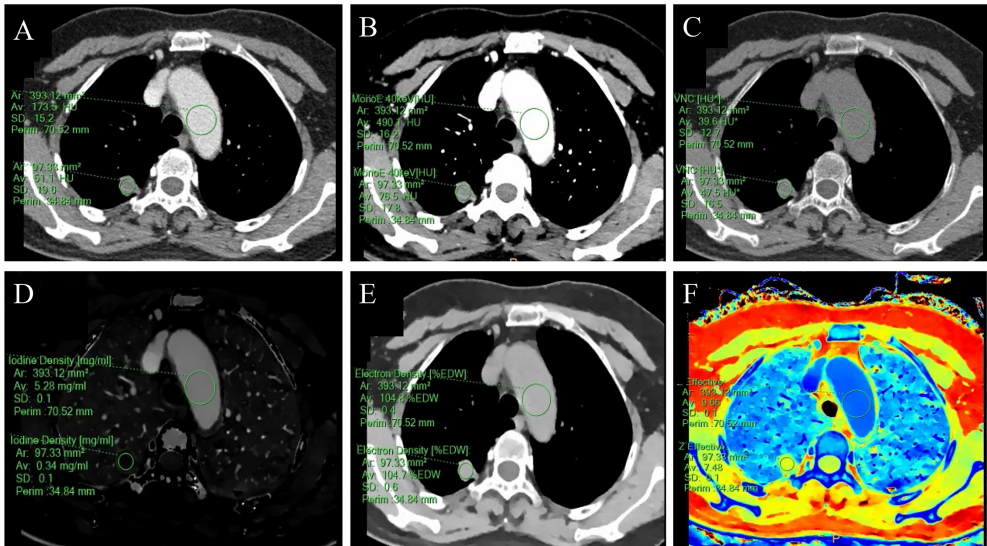


FIGURE 2 A 59-year-old female with a solid SPN in the posterior segment of the right upper lobe had pathologically confirmed hamartoma. The corresponding quantitative parameter maps of SDCT at 70keV, 40keV, virtual non-contrast (VNC), iodine density (IC), electron density (ED) and effective atomic number (Zeff) are shown in (A-F), respectively. The region of interest (ROI) outlined by the green circle reflects the corresponding SDCT quantitative parameters of the lesion and the aorta at the same level. In this case: CT_{SPN-70keV} = 51.1 HU, CT_{SPN-40keV} = 76.5 HU, CT_{SPN-VNC} = 47.5 HU, IC_{SPN} = 0.34 mg/ml, ED_{SPN} = 104.7% EDW, Z_{eff-SPN} = 7.48, CT_{aorta-70keV} = 173.5 HU, CT_{aorta-40keV} = 490.1 HU, CT_{aorta-VNC} = 39.6 HU, IC_{aorta} = 5.28 mg/ml, ED_{aorta} = 104.8% EDW, Z_{eff-aorta} = 9.66.

+Δ_{70keV}+NIC) was the largest (0.903) when the cut-off value was 0.728, the sensitivity, specificity, positive predictive rate, negative predictive rate and accuracy were 88.2%, 83.3%, 93.8%, 66.7% and 86.4%, respectively. The probability prediction model was logit(P) = -6.707 + 0.129size+0.059Δ_{70keV}+2.049NIC. The details are presented in Table 8 and Figure 5.

3.7 Assessment of inter-observer agreement

All SDCT parameters showed good inter-observer agreement for assessing SPNs, with ICC values between 0.811-0.997 (Table 9, Figure 6).

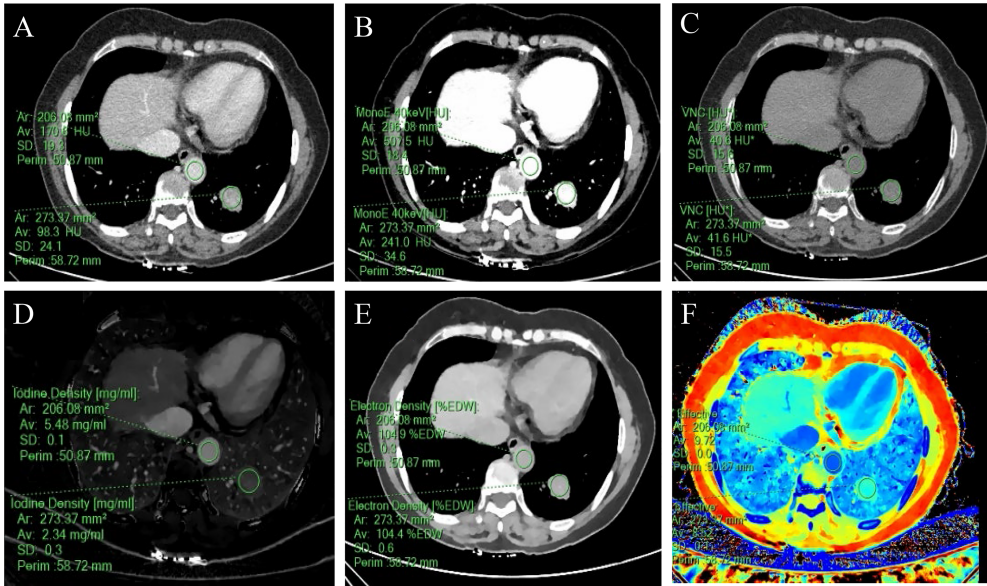


FIGURE 3
A 57-year-old woman with a solid SPN in the lateral basal segment of the left lower lobe had pathologically confirmed adenocarcinoma. The corresponding quantitative parameter maps of SDCT at 70keV, 40keV, virtual non-contrast (VNC), iodine density (IC), electron density (ED) and effective atomic number (Zeff) are shown in (A-F), respectively. The region of interest (ROI) outlined by the green circle reflects the corresponding SDCT quantitative parameters of the lesion and the aorta at the same level. In this case: $CT_{SPN-70keV} = 98.3$ HU, $CT_{SPN-40keV} = 241.0$ HU, $CT_{SPN-VNC} = 41.6$ HU, $IC_{SPN} = 2.34$ mg/ml, $ED_{SPN} = 104.4\%$ EDW, $Z_{eff-spn} = 8.52$, $CT_{aorta-70keV} = 170.6$ HU, $CT_{aorta-40keV} = 507.5$ HU, $CT_{aorta-VNC} = 40.6$ HU, $IC_{aorta} = 5.48$ mg/ml, $ED_{aorta} = 104.9\%$ EDW, $Z_{eff-aorta} = 9.72$.

TABLE 4 Differences in SDCT quantitative parameters among the sub-groups.

Parameters	Benign group (n=30)	Adenocarcinoma group (n=76)	Squamous cell carcinoma group(n=21)	P-value		
				Benign vs Adenocarcinoma	Benign Vs Squamous cell carcinoma	Adenocarcinoma vs Squamous cell carcinoma
SAR_{VNC}	0.707 ± 0.357	0.765 ± 0.242	0.844 ± 0.177	0.788	0.209	0.282
SAR_{40keV}	0.512 ± 0.198	0.852 ± 0.288	0.800 ± 0.207	<0.001*	<0.001*	0.739
SAR_{70keV}	0.240 ± 0.098	0.352 ± 0.110	0.344 ± 0.086	<0.001*	<0.001*	0.754
Δ_{40keV}	83.884 ± 43.997	150.456 ± 46.649	141.389 ± 47.793	<0.001*	<0.001*	0.824
Δ_{70keV}	21.944 ± 13.548	42.541 ± 13.827	39.600 ± 15.328	<0.001*	<0.001*	0.812
CER_{40keV}	3.262 ± 2.480	5.374 ± 3.703	4.076 ± 1.744	0.003*	0.434	0.075
CER_{70keV}	0.830 ± 0.699	1.528 ± 1.090	1.138 ± 0.545	0.001*	0.230	0.078
NEF_{40keV}	0.462 ± 0.212	0.894 ± 0.398	0.795 ± 0.249	<0.001*	0.001*	0.244
NEF_{70keV}	0.120 ± 0.067	0.254 ± 0.120	0.221 ± 0.073	<0.001*	0.001*	0.197
λ	1.990 ± 1.091	3.529 ± 1.101	3.320 ± 1.109	<0.001*	<0.001*	0.442
NIC	0.134 ± 0.064	0.264 ± 0.108	0.229 ± 0.075	<0.001*	<0.001*	0.264
NED	0.961 ± 0.089	0.980 ± 0.112	0.985 ± 0.009	0.572	0.404	0.191
NZ_{eff}	0.764 ± 0.040	0.809 ± 0.088	0.789 ± 0.528	0.005*	0.237	0.271

*p value<0.05.

4 Discussion

A multi-centre retrospective study by Tanner et al. reported that up to 35% of patients with benign pulmonary nodules received

surgical treatment, which not only increased the risk of complications and medical costs but also unnecessary psychological stress and social burden (20). Interestingly, our study suggests that SDCT quantitative parameters combined with

TABLE 5 Diagnostic efficacy of SDCT quantitative parameters in benign and malignant groups.

Parameters	AUC	Cut-off value	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)	P-value
SAR _{40keV}	0.848	0.617	79.4	76.7	92.0	52.3	78.2	<0.001*
SAR _{70keV}	0.783	0.331	53.9	90.0	94.8	36.5	61.7	<0.001*
Δ_{40keV}	0.836	113.217	75.5	80.0	92.8	49.0	75.9	<0.001*
Δ_{70keV}	0.846	33.000	71.6	86.7	94.8	47.3	74.4	<0.001*
CER _{40keV}	0.737	3.7354	57.8	80.0	89.4	34.8	61.7	<0.001*
CER _{70keV}	0.750	1.105	53.9	83.3	91.7	34.7	60.2	<0.001*
NEF _{40keV}	0.853	0.650	72.5	83.3	92.5	46.2	73.7	<0.001*
NEF _{70keV}	0.854	0.168	77.5	80.0	91.9	50.0	76.7	<0.001*
λ	0.831	2.662	76.5	80.0	92.9	50.0	76.7	<0.001*
NIC	0.869	0.165	87.3	76.7	92.7	63.9	84.2	<0.001*
NZ _{eff}	0.695	0.814	36.3	96.7	94.8	36.5	61.7	<0.001*

*p value<0.05.

TABLE 6 Diagnostic efficacy of SDCT quantitative parameters in benign and adenocarcinoma groups.

Parameters	AUC	Cut-off value	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)	P-value
SAR _{40keV}	0.842	0.656	75.0	80.0	89.1	54.8	74.8	<0.001*
SAR _{70keV}	0.771	0.331	50.0	90.0	92.7	41.5	60.7	<0.001*
Δ_{40keV}	0.849	113.217	78.9	80.0	90.9	60.0	78.5	<0.001*
Δ_{70keV}	0.857	33.0	75.0	86.7	93.4	57.8	77.6	<0.001*
CER _{40keV}	0.747	3.7354	61.8	80.0	87.0	44.2	65.4	<0.001*
CER _{70keV}	0.770	1.050	61.8	80.0	88.7	45.3	66.4	<0.001*
NEF _{40keV}	0.850	0.650	72.4	83.3	91.5	53.2	73.8	<0.001*
NEF _{70keV}	0.854	0.155	78.9	76.7	89.6	59.0	77.6	<0.001*
λ	0.842	2.662	78.9	80.0	90.9	60.0	78.5	<0.001*
NIC	0.876	0.158	92.1	73.3	89.7	78.6	86.0	<0.001*
NZ _{eff}	0.705	0.814	39.5	96.7	96.9	39.2	56.1	<0.001*

*p value<0.05.

morphological features are helpful in the diagnosis of solid SPNs with a diameter of 8–30 mm. Some SDCT parameters are also significant in differentiating the pathological subtypes of lung cancer presenting as solid SPNs. This study allowed patients to avoid invasive treatments and examinations. The combination of morphological and quantitative parameters has deepened the understanding of solid SPNs, which can better guide clinicians in developing the most appropriate treatment plan for patients with SPNs.

SPNs can be divided into solid SPN and sub-solid SPN based on the presence or absence of a ground-glass component, which includes pure ground-glass nodules (pGGN) and mixed ground-glass nodules (mGGN) (21). In this study, SDCT quantitative parameters were obtained by delineating the ROI of the solid

components of the lesion in the mediastinal window, while there were no solid or less solid components in the mediastinal window in the sub-solid SPN. Therefore, patients with sub-solid SPNs were not included in this study.

Radiological morphology features are important for the diagnosis and the differential diagnosis of SPNs (22, 23). In this study, CT findings of benign and malignant solid SPNs were analysed for 10 aspects, including location, size, lesion shape, lobulation, short spiculation, pleural indentation, vascular convergence sign, air bronchogram, vacuole sign, and calcification. There were significant differences in lesion size, shape, short spiculation sign, and vascular convergence sign between benign and malignant SPNs. Malignant SPNs are larger than benign SPNs because of its malignant characteristics and rapid

TABLE 7 Diagnostic efficacy of SDCT quantitative parameters in benign and squamous cell carcinoma groups.

Parameters	AUC	Cut-off value	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)	P-value
SAR _{40keV}	0.849	0.617	81.0	76.7	68.0	84.6	75.0	<0.001*
SAR _{70keV}	0.792	0.326	66.7	86.7	77.8	78.8	76.9	<0.001*
Δ_{40keV}	0.811	113.217	76.2	80.0	72.7	82.8	76.9	<0.001*
Δ_{70keV}	0.817	31.93	71.4	83.3	71.4	80.0	75.0	<0.001*
NEF _{40keV}	0.849	0.650	71.4	83.3	75.0	80.6	76.9	<0.001*
NEF _{70keV}	0.837	0.168	81.0	80.0	70.8	85.2	76.9	<0.001*
λ	0.806	2.662	76.2	80.0	72.7	82.8	76.9	<0.001*
NIC	0.832	0.165	81.0	76.7	70.8	85.2	76.9	<0.001*

*p value<0.05.

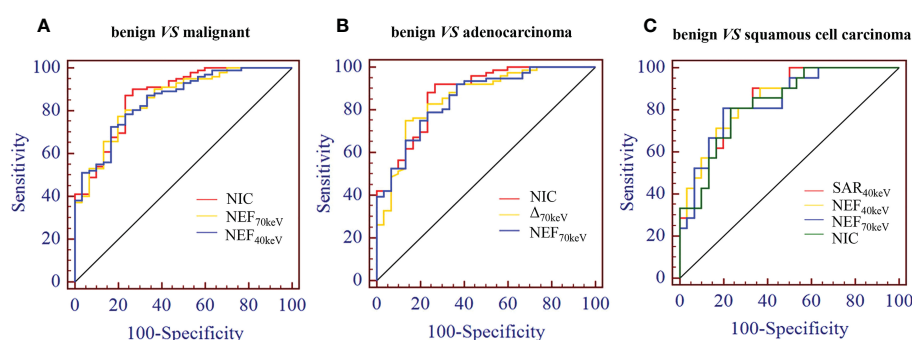


FIGURE 4

(A) Receiver operating characteristic curve (ROC) showed that NIC, NEF_{70keV} and NEF_{40keV} had the highest diagnostic efficiency in benign and malignant SPNs, and the AUC were 0.869, 0.854 and 0.853, respectively, of which NIC was the highest. (B) The three parameters with the highest diagnostic efficiency were NIC, Δ_{70keV} and NEF_{70keV}, and their AUC were 0.876, 0.857 and 0.854, respectively. NIC still had the highest detection efficiency. (C) SAR_{40keV}, NEF_{40keV} and NEF_{70keV} were the three parameters with the highest diagnostic efficiency to distinguish benign SPN from the squamous cell carcinoma group. The AUC were 0.849, 0.849 and 0.837, respectively. The diagnostic efficiency of NIC was 0.832, which was second only to the above three parameters.

growth rate. This study shows that lesion size is closely related and an independent risk factor for the prediction of benign and malignant SPNs, which is consistent with previous reports (24, 25). Owing to the heterogeneity of tumour cells, the growth pattern of lesions can be multi-polarised, which also leads to the irregular morphology of malignant tumours (26, 27). The short spiculation sign refers to the radial short thin-line shadow that extends from the edge of the mass to the surrounding lung parenchyma and is not connected with the adjacent pleura (28). Yi et al. reported that the malignant probability of SPNs with a spicular sign is 88%-94% (29). The sign of vascular convergence is also common in malignant SPNs. On the one hand, the cancer focus pulls adjacent blood vessels off the original track and releases a large amount of vascular endothelial growth factor that promotes the formation of new blood vessels, enhances tumour growth, and provides a path for metastasis (30). The lobulation sign, pleural indentation sign, air bronchogram sign, and bubble sign are also important in the diagnosis of malignant SPNs (31, 32). However, this study showed no significant difference between the benign and malignant groups. We hypothesised that since the average size of malignant lesions is

18.225 ± 5.689 mm, which is relatively small; the significance of the imaging parameters could not be achieved. Also, the number of included cases was insufficient.

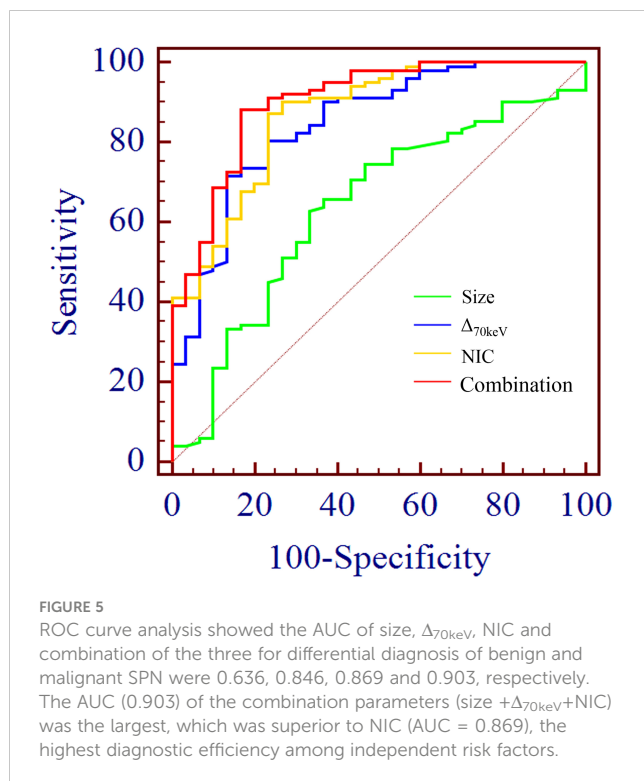
As a new CT technique, SDCT not only can obtain traditional CT images but also acquire a variety of spectral parameter images (such as VNC, virtual monoenergetic image (VMI), IC, Z_{eff} , Uric Acid, and Calcium Suppression). These spectral images can be used to reduce artefacts, improve image quality, reduce contrast agent dosage, and reduce the radiation dose. At the same time, they can also provide more valuable information for the detection of lesions, accurate measurement of lesion size, and differentiation of benign and malignant lesions (33–35).

The application of enhanced CT-derived parameters, CT enhancement value, and enhancement ratio could reduce the background influence caused by machine differences and individual differences and are considered to be effective tools for evaluating tumour angiogenesis. Moreover, CT enhancement values and enhancement ratios are related to microvascular and lymphatic invasion in tumours and can be used as surrogate markers for preoperative detection of lymphovascular invasion (19, 36, 37).

TABLE 8 Diagnostic efficacy of independent risk factors and the combination model.

Parameters	AUC	Cut-off value	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)	P-value
size	0.636	16.250	62.7	66.7	86.7	35.1	63.9	0.021*
$\Delta_{70\text{keV}}$	0.846	33.000	71.6	86.7	94.8	47.3	74.4	<0.001*
NIC	0.869	0.165	87.3	76.7	92.7	63.9	84.2	<0.001*
combination model	0.903	0.728	88.2	83.3	93.8	66.7	86.4	<0.001*

*p value<0.05.



Inspired by this, our study will focus, in the future, on the application of SDCT quantitative parameters and their derived parameters after normalisation to evaluate the benign and malignant nature of SPNs more objectively and accurately.

The study showed that between benign and malignant solid SPNs, except for SAR_{VNC} and NED ($P = 0.114, 0.208$), there were significant differences in other SDCT-related parameters ($\text{SAR}_{40\text{keV}}$, $\text{SAR}_{70\text{keV}}$, $\Delta_{40\text{keV}}$, $\Delta_{70\text{keV}}$, $\text{CER}_{40\text{keV}}$, $\text{CER}_{70\text{keV}}$, $\text{NEF}_{40\text{keV}}$, $\text{NEF}_{70\text{keV}}$, λ , NIC, NZ_{eff}) ($P < 0.05$). The parameters of malignancy were significantly higher which might be because malignant nodules produce a large number of angiogenic factors for tumour growth and stimulate the formation of more microvessels. The increased density of microvessels leads to increased capillary perfusion, and the contrast agent rapidly accumulates after enhancement, such that malignant nodules show a relatively strong contrast enhancement effect. This is consistent with previous related reports (38, 39). Z_{eff} is a quantitative index derived from the atomic number, which represents the composite atoms of compounds or mixtures of

various materials and characterizes the tissue composition. The NZ_{eff} value of the malignant SPNs in this study was higher than that of the benign group, which is also consistent with a recent study (40). A possible explanation could be that malignant lesions are densely packed tumour cells with a higher nuclear/cytoplasmic ratio. ROC curve analysis showed that NIC, $\text{NEF}_{70\text{keV}}$ and $\text{NEF}_{40\text{keV}}$ were the three best parameters for predicting the diagnosis of benign and malignant SPNs, respectively and the AUC of the NIC were 0.869, 0.854, and 0.853, respectively. This study also showed that NIC still had the best diagnostic performance in differentiating benign from adenocarcinoma SPNs (AUC=0.876) and that NIC also performed well in differentiating benign from squamous cell carcinoma (AUC=0.832). Iodine, as the main component of the contrast agent, directly reflects blood flow and distribution in the intravascular and extracellular spaces. The iodine concentration map is generally considered to have the potential to evaluate the number and blood flow of blood vessels supplying the pulmonary nodules. Using aortic iodine concentration as the standard parameter of NIC minimises the influence of haemodynamic factors on the absolute enhancement of lesions among different individuals, thereby increasing the comparability between different cases and making NIC more specific and stable than other iodine-related indicators. In this study, the NIC cut-off value for differentiating benign and malignant SPNs obtained during the 50 s of enhanced scanning was 0.165, which is in the range of 0.13 and 0.31 under the 25 s and 60 s enhanced scanning time as reported by Wen et al., which was consistent with previous reports (38, 41).

Binary logistic regression analysis, performed on the indicators and parameters that showed statistically significant differences between benign and malignant solid SPNs demonstrated that NIC, $\Delta_{70\text{keV}}$ and lesion size were independent risk factors for the prediction of malignant SPNs. When the three factors were combined for ROC curve analysis, the AUC for the differential diagnosis of benign and malignant SPNs was 0.903, the cut-off value was 0.728, and the sensitivity, specificity, positive prediction rate, negative prediction rate, and accuracy were 88.2%, 83.3%, 93.8%, 66.7%, and 86.4%, respectively. It was superior to the NIC (AUC=0.869), which had the highest diagnostic efficiency among the independent risk factors. The radiological image features reflect the macroscopic details of the lesion morphology, whereas the quantitative parameters of SDCT reflect microscopic details, such as blood flow and material composition. This study shows that the

TABLE 9 Repeatability of multi-parameters between observers.

Parameters	ICC (95% CI)	Parameters	ICC (95% CI)
SAR _{40keV}	0.849 (0.764-0.905)	SAR _{70keV}	0.910(0.858-0.944)
Δ_{40keV}	0.939(0.902-0.962)	Δ_{70keV}	0.957(0.931-0.974)
CER _{40keV}	0.820(0.722-0.886)	CER _{70keV}	0.845(0.759-0.902)
NEF _{40keV}	0.915 (0.866-0.947)	NEF _{70keV}	0.941 (0.906-0.964)
λ	0.916(0.867-0.948)	NIC	0.900(0.842-0.938)
NED	0.997(0.996-0.988)	NZ _{eff}	0.811(0.708-0.880)
SAR _{VNC}	0.898(0.839-0.936)		

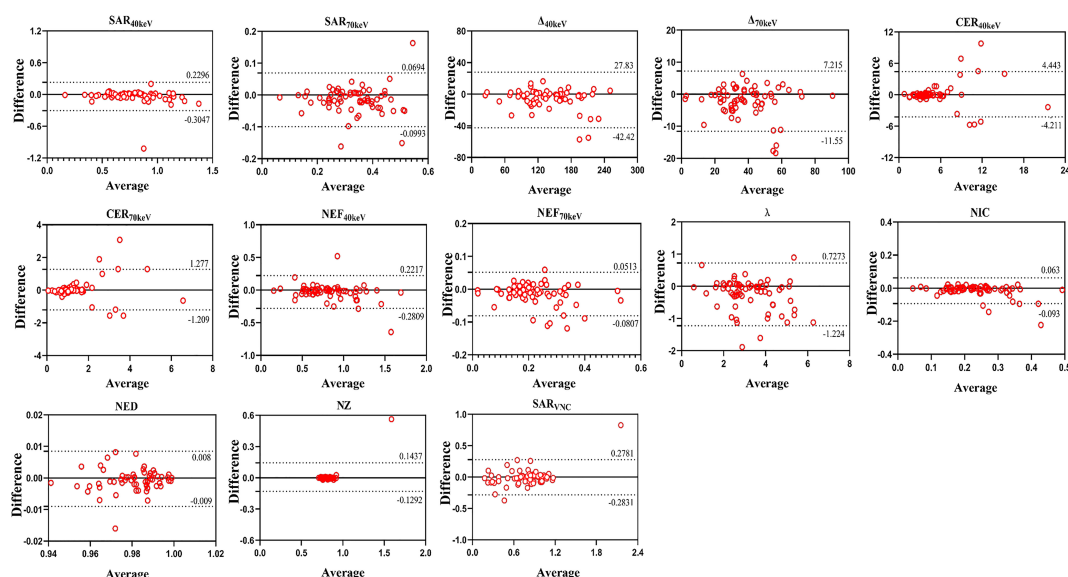


FIGURE 6
Bland-Altman plots of inter-observer differences in SDCT-related parameters.

combination of morphological signs and SDCT quantitative parameters could more accurately reveal the essential characteristics of solid SPNs, thereby further improving diagnostic efficiency, which has rarely been reported in previous studies.

This study has some limitations. First, this was a single-centre retrospective study with a relatively small sample size, and thus larger and multi-centre validation is needed. Second, to reduce the radiation dose, the corresponding SDCT quantitative parameters obtained from venous phase scan images were used in a single phase, although some studies have reported that some parameters (IC, λ , etc.) in the venous phase were better than those in the arterial phase (41, 42), and a multi-phase scanning study to further improve the integrity of the study will be performed, in the future. Third, the radiologists in our centre are experienced in the evaluation of benign and malignant pulmonary nodules, and most of the benign pulmonary nodules are followed up regularly; therefore, the number of benign cases undergoing surgery is small. In the subgroup analysis, the sample sizes of adenocarcinoma and

squamous cell carcinoma were small. In the future, a multi-centre collaboration will be considered to increase the sample size.

In conclusion, SDCT quantitative parameters are helpful for the differential diagnosis of benign and malignant solid SPNs, and NIC is superior to the other relevant quantitative parameters. When NIC is combined with lesion size and Δ_{70keV} value for comprehensive diagnosis, the efficacy power could be further improved.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and

institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

(I) Conception and design: MW, ZZ, WS; (II) Administrative support: MW, ZZ, WS; (III) Provision of study materials or patients: XX, KFL, YX; (IV) Collection and assembly of data: XX, KFL, KL, MW; (V) Data analysis and interpretation: XX, KFL, LZ; (VI) Manuscript writing: All authors; (VII) Final approval of the manuscript: All authors.

Funding

This work was supported by the Scientific Research Project of Jiangsu Provincial Health Commission [Grant Number Z2022075]

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancer in 185 countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
- Weir-McCall JR, Joyce S, Clegg A, MacKay JW, Baxter G, Dendl LM, et al. Dynamic contrast-enhanced computed tomography for the diagnosis of solitary pulmonary nodules: a systematic review and meta-analysis. *Eur Radiol* (2020) 30:3310–23. doi: 10.1007/s00330-020-06661-8
- He J, Li N, Chen WQ, Wu N, Shen HB, Jiang Y, et al. Consulting group of China guideline for the screening and early diagnosis and treatment of lung cancer; expert group of China guideline for the screening and early diagnosis and treatment of lung cancer; working group of China guideline for the screening and early diagnosis and treatment of lung cancer. [China guideline for the screening and early detection of lung cancer (2021, Beijing)]. *Zhonghua Zhong Liu Za Zhi* (2021) 43:243–68. doi: 10.11735/j.issn.1004-0242.2021.02.A001
- Tammemagi MC, Freedman MT, Pinsky PF, Oken MM, Hu P, Riley TL, et al. Prediction of true positive lung cancers in individuals with abnormal suspicious chest radiographs: a prostate, lung, colorectal, and ovarian cancer screening trial study. *J Thorac Oncol* (2009) 4:710–21. doi: 10.1097/JTO.0b013e31819e77ce
- Xu C, Hao K, Song Y, Yu L, Hou Z, Zhan P. Early diagnosis of solitary pulmonary nodules. *J Thorac Dis* (2013) 5:830–40. doi: 10.3978/j.issn.2072-1439.2013.11.19
- Sim YT, Goh YG, Dempsey MF, Han S, Poon FW. PET-CT evaluation of solitary pulmonary nodules: correlation with maximum standardized uptake value and pathology. *Lung* (2013) 191:625–32. doi: 10.1007/s00408-013-9500-6
- Zheng J, Ye X, Liu Y, Zhao Y, He M, Xiao H. The combination of CTCs and CEA can help guide the management of patients with SPNs suspected of being lung cancer. *BMC Cancer* (2020) 20:106. doi: 10.1186/s12885-020-6524-1
- Chu GCW, Lazare K, Sullivan F. Serum and blood based biomarkers for lung cancer screening: a systematic review. *BMC Cancer* (2018) 18:181. doi: 10.1186/s12885-018-4024-3
- Swensen SJ, Viggiano RW, Midthun DE, Muller NL, Sherrick A, Yamashita K, et al. Lung nodule enhancement at CT: multicenter study. *Radiology* (2000) 214:73–80. doi: 10.1148/radiology.214.1.r00ja1473
- Girvin F, Ko JP. Pulmonary nodules: detection, assessment, and CAD. *AJR Am J Roentgenol* (2008) 191:1057–69. doi: 10.2214/AJR.07.3472
- Deniffel D, Sauter A, Fingerle A, Rummeny EJ, Makowski MR, Pfeiffer D. Improved differentiation between primary lung cancer and pulmonary metastasis by combining dual-energy CT-derived biomarkers with conventional CT attenuation. *Eur Radiol* (2021) 31:1002–10. doi: 10.1007/s00330-020-07195-9
- McCollough CH, Leng S, Yu L, Fletcher JG. Dual- and multi-energy CT: principles, technical approaches, and clinical applications. *Radiology* (2015) 276:637–53. doi: 10.1148/radiol.2015142631
- Azour L, Ko JP, O'Donnell T, Patel N, Bhattacharji P, Moore WH. Combined whole-lesion radiomic and iodine analysis for differentiation of pulmonary tumors. *Sci Rep* (2022) 12:11813. doi: 10.1038/s41598-022-15351-y
- Yang X, Hu H, Zhang F, Li D, Yang Z, Shi G, et al. Preoperative prediction of the aggressiveness of oral tongue squamous cell carcinoma with quantitative parameters from dual-energy computed tomography. *Front Oncol* (2022) 12:904471. doi: 10.3389/fonc.2022.904471
- Chen W, Ye Y, Zhang D, Mao L, Guo L, Zhang H, et al. Utility of dual-layer spectral-detector CT imaging for predicting pathological tumor stages and histologic grades of colorectal adenocarcinoma. *Front Oncol* (2022) 12:1002592. doi: 10.3389/fonc.2022.1002592
- Wang X, Liu D, Zeng X, Jiang S, Li L, Yu T, et al. Dual-energy CT quantitative parameters for the differentiation of benign from malignant lesions and the prediction of histopathological and molecular subtypes in breast cancer. *Quant Imaging Med Surg* (2021) 11:1946–57. doi: 10.21037/qims-20-825
- Zou Y, Sun S, Liu Q, Liu J, Shi Y, Sun F, et al. A new prediction model for lateral cervical lymph node metastasis in patients with papillary thyroid carcinoma: based on dual-energy CT. *Eur J Radiol* (2021) 145:110060. doi: 10.1016/j.ejrad.2021.110060
- Huang S, Meng H, Cen R, Ni Z, Li X, Suwal S, et al. Use quantitative parameters in spectral computed tomography for the differential diagnosis of metastatic mediastinal lymph nodes in lung cancer patients. *J Thorac Dis* (2021) 13:4703–13. doi: 10.21037/jtd-21-385
- Ma Z, Liang C, Huang Y, He L, Liang C, Chen X, et al. Can lymphovascular invasion be predicted by preoperative multiphase dynamic CT in patients with advanced gastric cancer? *Eur Radiol* (2017) 27:3383–91. doi: 10.1007/s00330-016-4695-6
- Tanner NT, Aggarwal J, Gould MK, Kearney P, Diette G, Vachani A, et al. Management of pulmonary nodules by community pulmonologists: a multicenter observational study. *Chest* (2015) 148:1405–14. doi: 10.1378/chest.15-0630
- Nakata M, Saeki H, Takata I, Segawa Y, Mogami H, Mandai K, et al. Focal ground-glass opacity detected by low-dose helical CT. *Chest* (2002) 121:1464–7. doi: 10.1378/chest.121.5.1464
- Snoeckx A, Reyntjens P, Desbuquoit D, Spinhoven MJ, Van Schil PE, van Meerbeek JP, et al. Evaluation of the solitary pulmonary nodule: size matters, but do not ignore the power of morphology. *Insights Imaging* (2018) 9:73–86. doi: 10.1007/s13244-017-0581-2
- Erasmus JJ, Connolly JE, McAdams HP, Roggli VL. Solitary pulmonary nodules: part i. morphologic evaluation for differentiation of benign and malignant lesions. *Radiographics* (2000) 20:43–58. doi: 10.1148/radiographics.20.1.g00ja0343
- Yang L, Zhang Q, Bai L, Li TY, He C, Ma QL, et al. Assessment of the cancer risk factors of solitary pulmonary nodules. *Oncotarget* (2017) 8:29318–27. doi: 10.18632/oncotarget.16426
- Shi CZ, Zhao Q, Luo LP, He JX. Size of solitary pulmonary nodule was the risk factor of malignancy. *J Thorac Dis* (2014) 6:668–76. doi: 10.3978/j.issn.2072-1439.2014.06.22
- MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images:

and Nanjing Medical University Imaging Elite Talents Project [Grant Number 320.1140.2022.0510.005].

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

from the fleischner society 2017. *Radiology* (2017) 284:228–43. doi: 10.1148/radiol.2017161659

27. Choromańska A, Macura KJ. Evaluation of solitary pulmonary nodule detected during computed tomography examination. *Pol J Radiol* (2012) 77:22–34. doi: 10.12659/PJR.882967

28. Ferreira JR Jr, Oliveira MC, de Azevedo-Marques PM. Characterization of pulmonary nodules based on features of margin sharpness and texture. *J Digit Imaging* (2018) 31:451–63. doi: 10.1007/s10278-017-0029-8

29. Yi CA, Lee KS, Kim EA, Han J, Kim H, Kwon OJ, et al. Solitary pulmonary nodules: dynamic enhanced multi-detector row CT study and comparison with vascular endothelial growth factor and microvessel density. *Radiology* (2004) 233:191–9. doi: 10.1148/radiol.2331031535

30. Birau A, Ceausu RA, Cimpean AM, Gaje P, Raica M, Olariu T. Assessment of angiogenesis reveals blood vessel heterogeneity in lung carcinoma. *Oncol Lett* (2012) 4:1183–6. doi: 10.3892/ol.2012.893

31. Liu Y, Balagurunathan Y, Atwater T, Antic S, Li Q, Walker RC, et al. Radiological image traits predictive of cancer status in pulmonary nodules. *Clin Cancer Res* (2017) 23:1442–9. doi: 10.1158/1078-0432.CCR-15-3102

32. Harders SW, Madsen HH, Rasmussen TR, Hager H, Rasmussen F. High resolution spiral CT for determining the malignant potential of solitary pulmonary nodules: refining and testing the test. *Acta Radiol* (2011) 52:401–9. doi: 10.1258/ar.2011.100377

33. Lennartz S, Le Blanc M, Zopfs D, Große Hokamp N, Abdullayev N, Laukamp KR, et al. Dual-energy CT-derived iodine maps: use in assessing pleural carcinomatosis. *Radiology* (2019) 290:796–804. doi: 10.1148/radiol.2018181567

34. Hsu CC, Jeavon C, Fomin I, Du L, Buchan C, Watkins TW, et al. Dual-layer spectral CT imaging of upper aerodigestive tract cancer: analysis of spectral imaging parameters and impact on tumor staging. *AJNR Am J Neuroradiol* (2021) 42:1683–9. doi: 10.3174/ajnr.A7239

35. Daoud B, Cazejust J, Tavolaro S, Durand S, Pommier R, Hamrouni A, et al. Could spectral CT have a potential benefit in coronavirus disease (COVID-19)? *AJR Am J Roentgenol* (2021) 216:349–54. doi: 10.2214/AJR.20.23546

36. Yin XD, Huang WB, Lu CY, Zhang L, Wang LW, Xie GH. A preliminary study on correlations of triple-phase multi-slice CT scan with histological differentiation and intratumoral microvascular/lymphatic invasion in gastric cancer. *Chin Med J (Engl)* (2011) 124:347–51. doi: 10.3760/cma.j.issn.0366-6999.2011.03.005

37. Li Y, Su H, Yang L, Yue M, Wang M, Gu X, et al. Can lymphovascular invasion be predicted by contrast-enhanced CT imaging features in patients with esophageal squamous cell carcinoma? a preliminary retrospective study. *BMC Med Imaging* (2022) 22:93. doi: 10.1186/s12880-022-00804-7

38. Wen Q, Yue Y, Shang J, Lu X, Gao L, Hou Y. The application of dual-layer spectral detector computed tomography in solitary pulmonary nodule identification. *Quant Imaging Med Surg* (2021) 11:521–32. doi: 10.21037/qims-20-2

39. Zegadlo A, Zabicka M, Kania-Pudlo M, Maliborski A, Rozyk A, Sosnicki W. Assessment of solitary pulmonary nodules based on virtual monochrome images and iodine-dependent images using a single-source dual-energy CT with fast kVp switching. *J Clin Med* (2020) 9:2514. doi: 10.3390/jcm9082514

40. Shen H, Yuan X, Liu D, Huang Y, Wang Y, Jiang S, et al. Multiparametric dual-energy CT for distinguishing nasopharyngeal carcinoma from nasopharyngeal lymphoma. *Eur J Radiol* (2021) 136:109532. doi: 10.1016/j.ejrad.2021.109532

41. Zhang Y, Cheng J, Hua X, Yu M, Xu C, Zhang F, et al. Can spectral CT imaging improve the differentiation between malignant and benign solitary pulmonary nodules? *PloS One* (2016) 11:e0147537. doi: 10.1371/journal.pone.0147537

42. Hou WS, Wu HW, Yin Y, Cheng JJ, Zhang Q, Xu JR. Differentiation of lung cancers from inflammatory masses with dual-energy spectral CT imaging. *Acad Radiol* (2015) 22:337–44. doi: 10.1016/j.acra.2014.10.004

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to improve diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

