



Cancer Screening and Detection: Can AI Change The Game?

Authors: Helena Bradbury, EMJ, London, UK

Citation: EMJ Radiol. 2024;5[1]:12-14.
<https://doi.org/10.33590/emjradiol/LZIE6219>.



"We recognise that cancer detection is one of the main pillars of how to take care of the population today," remarked Luis Marti-Bonmati, Le Fe Polytechnic and University Hospital, Valencia, Spain, who chaired a session at the European Congress of Radiology (ECR). In a timely conversation, the role of artificial intelligence (AI) in the screening, early detection, and depiction of tumours was explored at the annual ECR congress, which took place in Vienna, Austria, from the 28th February–3rd March 2024.

AI AND BREAST CANCER SCREENING

Sarah J. Vinnicombe, Cheltenham General Hospital, UK, opened her presentation citing a particularly concerning statistic; it is estimated that by 2027, there will be a 40% reduction in consultant breast radiologists within the UK. AI could alleviate this potential workforce crisis by shortening reading time, improving workflow efficacy, and thus reducing radiologists' workload.

The traditional cancer screening workflow involves consultation between two readers, followed by arbitration. Within this process, AI could replace the second reader, acting as an aid for both radiologists, or as a pre-screening triage tool. Despite the improved resolution and image quality digital breast tomosynthesis offers over traditional mammograms, it takes almost twice as long, slowing down the efficiency of radiologists' workflow. Referencing a 2022 study, Vinnicombe proposed that AI could act as complimentary tool to digital breast tomosynthesis, with 17–91% of digital mammogram scans being able to be read by AI alone, missing only 0–7% of cancer cases.

Interval cancer is defined as breast cancer detected during the 3 years after a normal result, and before the next screening appointment. Characteristically, interval cancers are aggressive and are associated with a poor prognosis. Drawing on novel research, Vinnicombe stated that AI flags 20–50% of interval cancers at the prior screen, which were incorrectly deemed negative by human readers. Operating at a 99% specificity, a 2022 study also concluded that AI could correctly localise 27.5% of false negatives, and 12.0% of cases with minimal signs, at the prior screen. The ability of machine learning to not only localise potential tumours, but also correct human misreading is highly significant, elevating the predictive properties of screening.

In her concluding remarks, Vinnicombe detailed the current barriers and facilitators in breast cancer screening. According to a review, which analysed 107 papers looking at the implementation of AI in clinical radiology, the common limitations are data size, variability, quality, model transparency, and meaningful clinical evaluation. Conversely, the majority of papers consistently concluded that AI mainly aids in diagnostic performance and clinical workflow.

"It is estimated that by 2027 there will be a 40% reduction in consultant breast radiologists within the UK."



AI AND LUNG CANCER SCREENING

Bram Van Ginneken, Radboudumc, Nijmegen, the Netherlands, discussed the current challenges in lung cancer screening, notably the occurrence of false positives, leading to overdiagnosis and overtreatment. Like Vinnicombe, Ginneken too was optimistic that AI could help minimise false readings, workload, and overall expenditure.

Several trials have demonstrated the use of low-dose chest CT as a diagnostic tool for lung cancer, namely the NLST, and more recently, the Dutch-Belgian NELSON trial. Ginneken summarised a study, assessing the performance of a computer aided detection (CAD) algorithm, to recognise abnormal nodules, and classify their malignancy risk based on volume. Unsurprisingly, the average reading time per scan was less than 1 second, compared to 60 seconds for a radiologist, and CAD successfully matched the malignancy risk of 70% of scans to the recommended NELSON criteria.

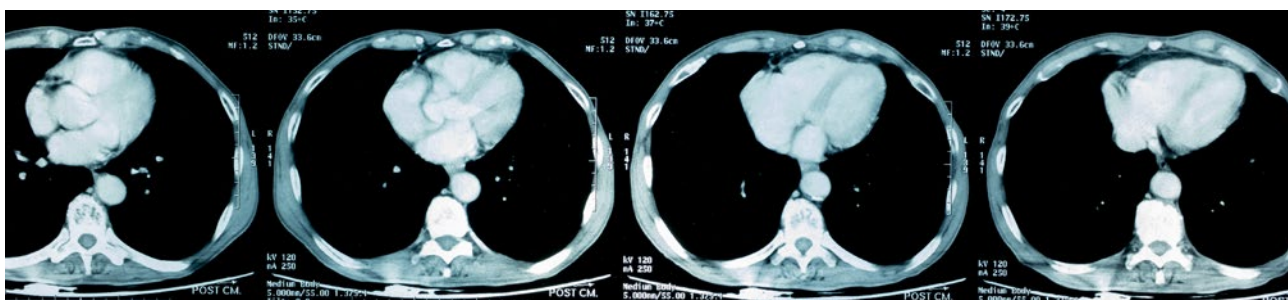
Ginneken emphasised the variability in performance between radiologists, but even between scans of an individual, and explained how AI offers, in comparison, consistent high performance. Looking to the future, he advocated for greater responsibility and tasks to be assigned to AI, allowing it to detect abnormalities across entire scans, and training it to identify rarer manifestations of the disease.

AI AND PANCREATIC CANCER

Vincenza Granata, Istituto Nazionale Tumori di Napoli, IRCCS G. Pascale, Naples, Italy, opened her presentation sharing the startling 5-year survival rate for pancreatic cancer in the USA (12%). With this statistic increasing to 44% with early detection, she stressed the importance of early intervention and screening techniques. The current guidelines, provided by the International Cancer of the Pancreas Screening (CAPS) consortium, recommend screening to start at 50–55 years for those who meet the familial risk criteria, and 40 years for patients with Peutz–Jeghers syndrome, a familial atypical mole, or melanoma syndrome.

"As the third leading cause of cancer deaths, PDAC is a significant global health concern."

Granata addressed several research initiatives set up to train and validate the use of deep learning models in the detection of pancreatic cancer, specifically pancreatic ductal adenocarcinoma (PDAC). As the third leading cause of cancer deaths, PDAC is a significant global health concern, with less than 20% of patients eligible for surgery at the time of diagnosis. The Felix project, for instance, is a multidisciplinary research collaboration, funded by the Lustgarten Foundation, New York, USA,



comprised of experts in medical imaging, pathology, cancer research, and computer science. Conducted at Johns Hopkins Hospital, Baltimore, Maryland, USA, it assessed the specificity and sensitivity of AI PDAC detection from CT scans. The preliminary results of this project, using 156 PDAC and 300 normal cases, were highly promising, with AI yielding a 94.1% sensitivity and 98.5% specificity for PDAC detection. Additionally, the artificial neural network (ANN) was developed, trained, and subsequently tested using the health data of 800,114 respondents, captured in the National Health Interview Survey (NHIS) and Pancreatic Lung, Colorectal and Ovarian Cancer (PLCO) datasets. Interestingly, ANN exhibited exceptional sensitivity and specificity, 87.3 and 80.3, respectively, with an area under the receiver operating curve of 0.85.

In her closing remarks, Grenata drew attention to several barriers preventing full integration of AI into clinical practice. Firstly, the training and validation of these models requires large datasets and multicentre studies, across various institutions and populations, to prevent opportunistic bias. She explained that the accuracy of AI detection is solely dependent on the image quality of CT scans, a factor that can be variable, especially in heterogeneous, multicentre datasets. Finally, she highlighted the extensive collaboration between several specialists, such as radiation oncologists, surgeons, and researchers, as well as policy makers, before the implementation of these predictive tools in patient care, and pancreatic cancer detection, can be a reality.

AI AND PROSTATE CANCER DETECTION

“MRI offers high sensitivity, approximately 91%, but lower specificity (37%) and moderate

reproducibility,” stated Olivier Rouvière, Centre Hospitalier Universitaire de Lyon, France, shifting the focus to the use of AI in prostate cancer detection. He outlined two fully automated systems, CADe and CADx, utilised in the detection and diagnosis of prostate cancer, respectively. Whilst CADe analyses MRI scans and highlights possible lesions, CADx quantifies the degree of suspicion of said lesions.

Although current research indicates exceptional detection capabilities in these systems, Rouvière cautiously noted some considerations when reading said literature. He put into question the definition of ‘external cohort’, a term mentioned frequently in validation studies. He explained that if the cohort AI is being tested on is too similar to the training dataset, the model will undoubtedly perform well, invalidating the study, and leading to opportunistic bias. As a combative effort, he called for large-scale external validation studies on multicentre, multivendor, multi-scanner, multiprotocol cohorts, to ensure thorough testing of the robustness of algorithms for prostate cancer detection. Finally, he touched on the potential shifts in sensitivity/specificity balance of predefined diagnostic thresholds.

CONCLUSION

With the ever-growing demand on diagnostic services, coupled with the deficit in clinical radiologists, the emergence of AI could not have come at a better time. By training computer models to detect abnormalities in scans, tasks traditionally performed by radiologists can be shared out, alleviating the workforce crisis, and revolutionising detection technology. However, as alluded to by the experts, cancer diagnosis is a complex process, and before AI can be routinely implemented into clinical practice, it must be thoroughly validated first. ●