EDITORIAL COMMENT

Is the future of breast imaging with AI?

Michael Fuchsjäger[1] (ORCID)

Where do we stand with computer aid in breast imaging?

Computers have been companions of breast imagers since the late 1990s, when the first CAD (computer-assisted detection) systems for—then conventional—mammography were introduced into clinical practice. The rationale was to help detect potential breast carcinomas that might otherwise be overlooked due to, for example, dense parenchyma. As with many new technologies, initial studies were encouraging, reporting higher sensitivity for the combination of mammography and CAD than for mammography alone [1]. However, over time, based on the aggregation of evidence from high-volume studies, it was determined that the use of CAD during the interpretation of mammography screening improved neither overall accuracy [2] nor any other relevant measure in the general patient population or in any subgroup of women [3].

Looking back, traditional CAD systems ultimately failed because they were feature-driven—using patterns perceived by humans (density, heterogeneity, shape, margins) to determine the presence or absence of specific imaging features—and therefore could not exceed the maximum accuracy of human breast imagers [4].

Where does AI come in?

When AI entered the scene some years ago, the radiology community was in an uproar, because the implications and consequences of AI were unclear, and radiologists were unsettled by newspaper reports of their imminent replacement by machines. Now, a few years later, the dust has settled, and AI's potential in radiology is being evaluated on many fronts. The absolute and quick disruption predicted by many has not happened thus far, and so the possible integration of AI and deep learning into clinical image interpretation appears likely to be evolutionary rather than revolutionary in nature.

A universal need to improve breast cancer screening outcomes, recent advances in computing power, and the sheer vastness of digital data available today (an estimated 100,000,000 screening mammograms performed around the world each year), which has led to concerns about radiologists' workloads, have paved the way for AI in breast imaging.

A class of machine learning methods has emerged that, unlike CAD, makes it possible to automatically discover the best features for any given task (i.e., through unsupervised learning), without requiring human feature engineering [4]. Digital images can be analyzed down to the individual pixel, providing virtually infinite variables. This has opened the door for the development of complex AI algorithms that could exceed human performance [5]. In short, AI detects relevant features that are imperceptible to the human eye.

A review of the first studies of AI for breast cancer detection showed that these were predominantly retrospective studies based on relatively small and narrowly selected (i.e., cancer-enriched) imaging data sets as well as heterogeneous AI techniques. Most studies validated developed models, and rather few tested models on independent datasets [6].

What does the current study tell us about AI in breast imaging?

Currently, AI seems to be most promising in very specific fields or niches like screening mammography. Therefore, breast imagers have high expectations that the main issues, namely missed cancers and false-positives resulting in overdiagnosis and potential overtreatment, can now be tackled. Recently, the working group around Rodriguez-Ruiz found that breast radiologists achieved better diagnostic performance

---

This comment refers to the article available at https://doi.org/10.1007/s00330-019-06186-9.

✉ Michael Fuchsjäger
michael.fuchsjaeger@medunigraz.at

1 Clinical Division of General Radiology, Department of Radiology, Medical University of Graz, Auenbruggerplatz 9, 8036 Graz, Austria

when using an AI algorithm for mammography as a decision support tool than when reading unaided [7]. In another recent study, the same group reported that this very AI algorithm, used as a stand-alone interpretation tool, achieved a cancer detection rate comparable with that of an average breast radiologist reading [8].

Another fundamental issue for radiologists—the ever-increasing workload—is addressed in a study by Rodriguez-Ruiz and coworkers in this issue of *European Radiology* [9]. A commercially available AI system (Transpara, version 1.4.0; Screenpoint Medical BV) stratified 2654 mammography examinations on a scale from 1 to 10 with regard to the likelihood of cancer presence. Score 1 indicated the lowest probability of cancer, score 10 the highest. In short, the study found that workload could be reduced by 47% by reading mammograms with scores from 6 to 10 only. Cases with scores of 1 to 5 were automatically assigned a normal report, resulting in a 7% rate of missed cancers. The authors concluded that these cancers might have been missed by radiologists anyhow due to their low mammographic visibility.

This is a highly interesting approach that needs closer examination. The algorithm could be used to reduce workload by excluding cases with a low probability of breast cancer from being read by radiologists at all. However, this very algorithm could also be used to select cases needing double reading versus cases needing single reading only; in this scenario, the workload reduction (of approximately 24%) would not be as dramatic, but a radiologist would read every single mammogram, which is one of the most critical aspects here.

As with previous screening mammography AI studies, the limitations of this study are its retrospective design and its reliance on a rather small, cancer-enriched sample size, without specific histopathology information and with single reading only.

In summary, the current study presents a new and possibly effective AI-based strategy to reduce the reading workload in mammography breast cancer screening programs without decreasing the sensitivity achieved by the average radiologist.

Where do we go from here?

Further down this road of excluding low probability mammography studies from human reading, we will have to ask our patients if they feel comfortable. And even if they have absolute trust in AI correctly interpreting their mammograms, do we?

Proper independent validation, as well as large prospective studies representing real-world screening scenarios to establish an evidence base, will be needed to evaluate developed AI models. AI algorithms should be expanded to cover other fields of breast imaging with huge amounts of data such as digital breast tomosynthesis, MRI, and automated 3D US.

Moreover, future research will have to deal with the acceptability of using AI in breast cancer screening services and the many ethical, social, and legal implications of their use [6]. Computers cannot be held accountable, but radiologists applying AI algorithms for reporting will be.

AI brings new opportunities—just imagine an algorithm identifying invasive breast cancers that will become clinically relevant only and omitting irrelevant ones from consideration. However, the fundamentals of clinical reality are unchanged; true success is defined by patient well-being and survival, not a better image quality or an improvement in statistical results alone [4].

AI will help make our lives easier by improving quality. With AI, radiologists will be able to generate further value as set forth by the framework of value-based radiology [10]. At the end of the day, we have to assure that the right examination is performed at the right time for the right patient. We have a good reason to be excited about AI as another asset that will enhance our ability to provide the best possible care for our patients.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Michael Fuchsjäger.

Conflict of interest The author of this manuscript declares no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent No informed consent needed.

Ethical approval Institutional Review Board approval was not required because this is an editorial comment.

Methodology
• Editorial comment

References

1. Funovics M, Schamp S, Helbich TH et al (2001) Evaluation of a computer-assisted diagnosis system in breast carcinoma. Rofo 173: 218–223
2. Fenton JJ, Taplin SH, Carney PA et al (2007) Influence of computer-aided detection on performance of screening mammography. N Engl J Med 356:1399–1409
3. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL (2015) Diagnostic accuracy of digital screening

mammography with and without computer-aided detection. JAMA Intern Med 175:1828–1837

4. Savadjiev P, Chong J, Dohan A et al (2019) Demystification of AI-driven medical image interpretation: past, present and future. Eur Radiol 29:1616–1624

5. Trister AD, Buist DSM, Lee CI (2017) Will machine learning tip the balance in breast cancer screening? JAMA Oncol 3:1463–1464

6. Houssami N, Kirkpatrick-Jones G, Noguchi N, Lee CI (2019) Artificial Intelligence (AI) for the early detection of breast cancer: a scoping review to assess AI's potential in breast screening practice. Expert Rev Med Devices 16:351–362

7. Rodríguez-Ruiz A, Krupinski E, Mordang JJ et al (2019) Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology 290:305–314

8. Rodriguez-Ruiz A, Lång K, Gubern-Merida A et al (2019b) Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst. https://doi.org/10.1093/jnci/djy222

9. Rodriguez-Ruiz A, Lång K, Gubern-Merida A (2019a) Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. Eur Radiol. https://doi.org/10.1007/s00330-019-06186-9

10. European Society of Radiology (ESR) (2017) ESR concept paper on value-based radiology. Insights Imaging 8:447–454