AI+Radiologist: Improved Screening Interpretation

Quality of breast cancer screenings can be improved through utilisation of AI algorithms, but it is necessary for human radiologists to participate in the evaluation process.

You might also like: **Suppressing Breast Cancer Metastasis**

These are the findings of a study (Schaffter et al. 2020) designed to compare performance of deep learning algorithms and radiologists in screening mammography interpretation and see whether AI can overcome human limitations. The study was based on the Digital Mammography (DM) Dialogue on Reverse Engineering Assessment and Methods (DREAM) challenge, a year-long international competition held from September 2016 to November 2017.

More than 1,100 participants comprising 126 teams from 44 countries participated. They used 144,231 screening mammograms from 85,580 women from the U.S. to train and validate their algorithms and independently confirmed their findings with a second dataset of over 166,578 screening mammograms from 68,008 women from Sweden.

The researchers set two 'challenges' for algorithms, using images alone (challenge 1) and combining images with previous examinations (if available) and clinical and demographic risk factor data (challenge 2). The resulting score translated to cancer yes/no within 12 months from screening (952 cancer positive in the U.S. sample, 780 cancer positive in the Swedish sample).

To assess algorithm accuracy for breast cancer detection the researchers used an area under the curve, and algorithm specificity was compared with radiologists' specificity. Radiologists' sensitivity was set at 85.9% (the U.S.) and 83.9% (Sweden).

You might also like: Deep Mining Reveals Omics Promise for Cancer Biomarkers

No single AI algorithm outperformed radiologists. For the top-performing algorithm, an area under the curve was 0.858 for the U.S. and 0.903 for Sweden and, respectively, 66.2% and 81.2% specificity at the radiologists' sensitivity. This is lower than community radiologist benchmarks, 90.5% in the U.S. and 98.5% in Sweden.
However, when combined with single-radiologist assessment, the algorithm achieved an area under the curve of 0.942 and a significantly improved specificity (92.0%) at the same sensitivity.

Overall accuracy of mammography screening interpretation could, therefore, be increased by integrating AI in single-radiologist settings and complementing human interpretation. Consequently, it may help to both cut healthcare system expenditures and increase efficiency of population-based screening programmes.

It is noted that sensitive patient information was protected through the model-to-data approach. It prevents the distribution of data to participants and mitigates the risk of sensitive patient data being released.

According to Dr Diana Buist of Kaiser Permanente Washington Health Research Institute who also co-authored the paper, this approach allowed participants to contribute innovations without receiving access to the underlying data. "Also, the inclusion of data from two different countries with differing mammography screening practices highlights important translational differences in how AI could be used in different populations," Dr Buist said in a statement.

The Digital Mammography DREAM Challenge was conducted by IBM Research, Sage Bionetworks, Kaiser Permanente Washington Health

Research Institute, and the University of Washington School of Medicine.

References

Schaffter T, Buist DSM, Lee CI, et al. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA Netw Open.* 2020;3(3):e200265. Available from https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2761795

Source: AI in Healthcare
Image credit: Pixabay

Published on : Wed, 4 Mar 2020