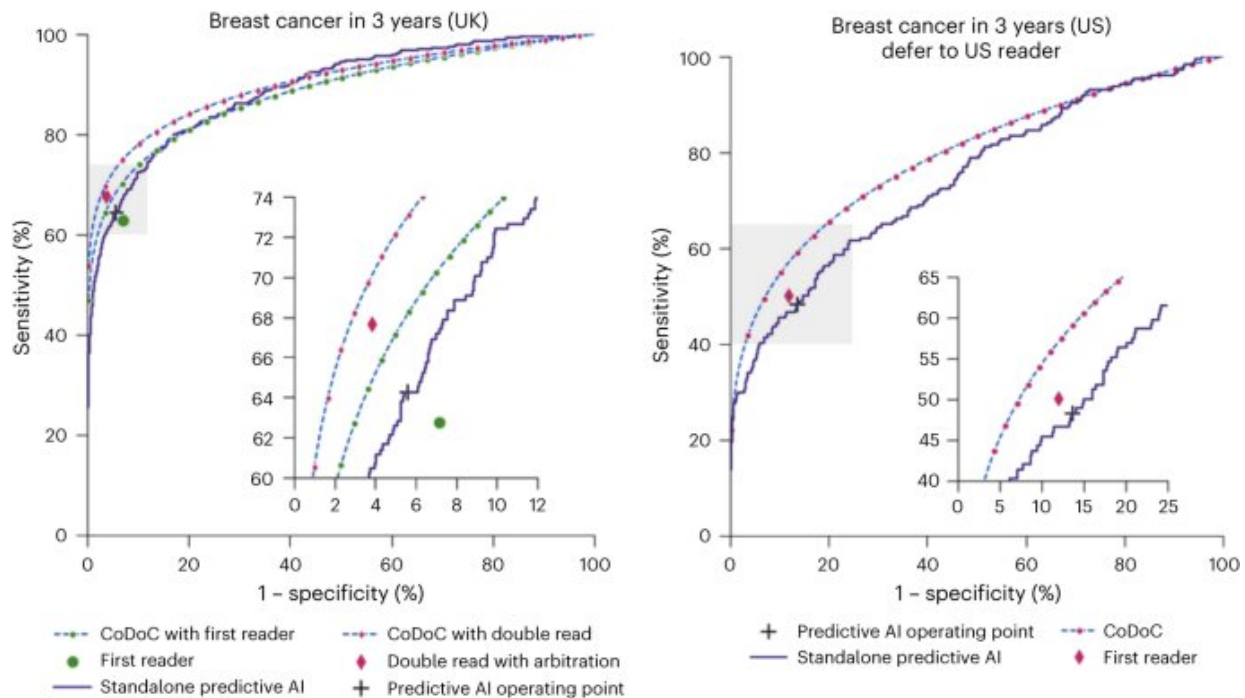# An AI system to figure out when to trust AI-based medical diagnoses

July 18 2023, by Bob Yirka

Performance of CoDoC in breast cancer prediction compared to that of a standalone predictive AI system and clinical readers. Credit: *Nature Medicine* (2023). DOI: 10.1038/s41591-023-02437-x

A team of AI and medical specialists working with or for Google Research and Google DeepMind, has developed an AI based system designed to judge the confidence level of existing AI systems used for analyzing medical scans as a means of improving analysis of diagnostic

tools, such as mammograms or chest X-rays.

In their paper, published in the journal *Nature Medicine,* the group describes how they built the system and how well it worked when tested. Fiona Gilbert, with the University of Cambridge's Clinical School of Medicine, has published a News & Views piece in the same journal issue outlining the work done by the team on this new effort.

Over the past several years, as AI applications have become more refined, the medical establishment has embraced the technology as a way to reduce the amount of work done by human radiologists while maintaining or improving quality. Currently, AI applications are used to analyze scans such as mammograms and X-rays looking for tumors in the breast or lungs.

Prior research has shown that the most reliable approach is to have both a human being and an AI app analyze the same scans. That way, fewer tumors are missed. Analysis by two radiologists gives nearly the same results.

In this new effort, the researchers sought to improve on this system by analyzing the results given by the AI systems working in such situations. To achieve that goal, they created a system called Complementarity-driven Deferral-to-Clinical Workflow (CoDoC). It has been designed to work alongside current AI systems that are already being used in the field and to use metrics already provided by such systems.

The new system's job is to analyze the results given by the diagnostic AI system and then to judge the degree of confidence in its results—and to give that information to human diagnosticians who provide the final diagnosis.

CoDoC is also an AI system. It was trained using gold-standard outcome

data obtained from deferral cases. The research team designed it with an improvement loop, which in theory should make it more accurate the more it is used. For now, it was designed to work with both a diagnostic AI system and a human radiologist.

In the end, CoDoC is used to help figure out which results are most likely correct, and that should result in improving accuracy. Testing of the system showed that to be the case. Scenarios where an AI system was used along with CoDoC and also a radiologist were found to be more reliable than either system alone.

More information: Krishnamurthy Dvijotham et al, Enhancing the reliability and accuracy of AI-enabled diagnosis via complementarity-driven deferral to clinicians, *Nature Medicine* (2023). DOI: 10.1038/s41591-023-02437-x

Fiona Gilbert, Balancing human and AI roles in clinical imaging, *Nature Medicine* (2023). DOI: 10.1038/s41591-023-02441-1

© 2023 Science X Network

Citation: An AI system to figure out when to trust AI-based medical diagnoses (2023, July 18) retrieved 12 May 2024 from https://medicalxpress.com/news/2023-07-ai-figure-ai-based-medical.html