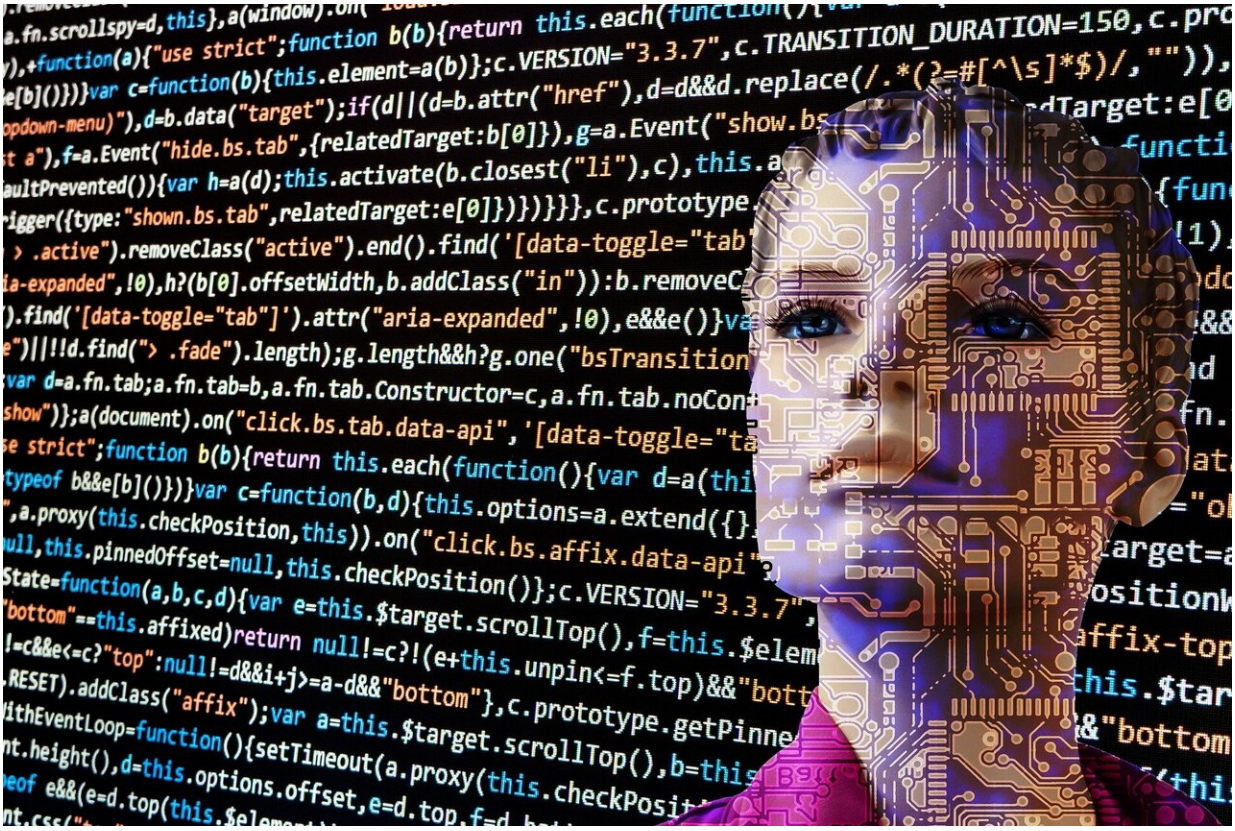
Study shows AI improves accuracy of skin cancer diagnoses

April 12 2024, by Krista Conger



Credit: CC0 Public Domain

A new study led by researchers at Stanford Medicine finds that computer algorithms powered by artificial intelligence based on deep learning can help health care practitioners to diagnose skin cancers more

accurately. Even dermatologists benefit from AI guidance, although their improvement is less than that seen for non-dermatologists.

"This is a clear demonstration of how AI can be used in collaboration with a physician to improve patient care," said professor of dermatology and of epidemiology Eleni Linos, MD. Linos leads the Stanford Center for Digital Health, which was launched to tackle some of the most pressing research questions at the intersection of technology and health by promoting collaboration between engineering, computer science, medicine and the humanities.

Linos, associate dean of research and the Ben Davenport and Lucy Zhang Professor in Medicine, is the senior author of the study, which was published in *npj Digital Medicine*. Postdoctoral scholar Jiyeong Kim, Ph.D., and visiting researcher Isabelle Krakowski, MD, are the lead authors of the research.

"Previous studies have focused on how AI performs when compared with physicians," Kim said. "Our study compared physicians working without AI assistance with physicians using AI when diagnosing skin cancers."

AI algorithms are increasingly used in clinical settings, including dermatology. They are created by feeding a computer hundreds of thousands or even millions of images of skin conditions labeled with information such as diagnosis and patient outcome.

Through a process called deep learning, the computer eventually learns to recognize telltale patterns in the images that correlate with specific skin diseases including cancers. Once trained, an algorithm written by the computer can be used to suggest possible diagnoses based on an image of a patient's skin that it has not been exposed to.

These diagnostic algorithms aren't used alone, however. They are overseen by clinicians who also assess the patient, come to their own conclusions about a patient's diagnosis and choose whether to accept the algorithm's suggestion.

An accuracy boost

Kim and Linos' team reviewed 12 studies detailing more than 67,000 evaluations of potential skin cancers by a variety of practitioners with and without AI assistance. They found that, overall, health care practitioners working without aid from artificial intelligence were able to accurately diagnose about 75% of people with skin cancer—a statistical measurement known as sensitivity. Conversely, the workers correctly diagnosed about 81.5% of people with cancer-like skin conditions but who did not have cancer—a companion measurement known as specificity.

Health care practitioners who used AI to guide their diagnoses did better. Their diagnoses were about 81.1% sensitive and 86.1% specific. The improvement may seem small, but the differences are critical for people told they don't have cancer, but do, or for those who do have cancer but are told they are healthy.

When the researchers split the health care practitioners by specialty or level of training, they saw that medical students, nurse practitioners and primary care doctors benefited the most from AI guidance—improving on average about 13 points in sensitivity and 11 points in specificity. Dermatologists and dermatology residents performed better overall, but the sensitivity and specificity of their diagnoses also improved with AI.

"I was surprised to see everyone's accuracy improve with AI assistance, regardless of their level of training," Linos said. "This makes me very optimistic about the use of AI in clinical care. Soon our patients will not

just be accepting, but expecting, that we use AI assistance to provide them with the best possible care."

Researchers at the Stanford Center for Digital Health, including Kim, are interested in learning more about the promise of and barriers to integrating AI-based tools into health care. In particular, they are planning to investigate how the perceptions and attitudes of physicians and patients to AI will influence its implementation.

"We want to better understand how humans interact with and use AI to make clinical decisions," Kim said.

Previous studies have indicated that a clinician's degree of confidence in their own clinical decision, the degree of confidence of the AI, and whether the clinician and the AI agree on the diagnosis all influence whether the clinician incorporates the algorithm's advice when making clinical decisions for a patient.

Medical specialties like dermatology and radiology, which rely heavily on images—visual inspection, pictures, X-rays, MRIs and CT scans, among others—for diagnoses are low-hanging fruit for computers that can pick out levels of detail beyond what a human eye (or brain) can reasonably process. But even other more symptom-based specialties, or prediction modeling, are likely to benefit from AI intervention, Linos and Kim feel. And it's not just patients who stand to benefit.

"If this technology can simultaneously improve a doctor's diagnostic accuracy and save them time, it's really a win-win. In addition to helping patients, it could help reduce physician burnout and improve the human interpersonal relationships between doctors and their patients," Linos said.

"I have no doubt that AI assistance will eventually be used in all medical

specialties. The key question is how we make sure it is used in a way that helps all patients regardless of their background and simultaneously supports physician well-being."

More information: Isabelle Krakowski et al, Human-AI interaction in skin cancer diagnosis: a systematic review and meta-analysis, *npj Digital Medicine* (2024). DOI: 10.1038/s41746-024-01031-w

Provided by Stanford University