The vast majority of AI models used in medicine today are "narrow specialists," trained to perform one or two tasks, such as scanning mammograms for signs of breast cancer or detecting lung disease on chest X-rays.

But the everyday practice of medicine involves an endless array of clinical scenarios, symptom presentations, possible diagnoses, and treatment conundrums. So, if AI is to deliver on its promise to reshape clinical care, it must reflect that complexity of medicine and do so with high fidelity, says Pranav Rajpurkar, assistant professor of biomedical informatics in the Blavatnik Institute at HMS.

Enter generalist medical AI, a more evolved form of machine learning capable of performing complex tasks in a wide range of scenarios.

Akin to general medicine physicians, Rajpurkar explained, generalist medical AI models can integrate multiple data types -- such as MRI scans, X-rays, blood test results, medical texts, and genomic testing -- to perform a range of tasks, from making complex diagnostic calls to supporting clinical decisions to choosing optimal treatment. And they can be deployed in a variety of settings, from the exam room to the hospital ward to the outpatient GI procedure suite to the cardiac operating room.

While the earliest versions of generalist medical AI have started to emerge, its true potential and depth of capabilities have yet to materialize.

"The rapidly evolving capabilities in the field of AI have completely redefined what we can do in the field of medical AI," writes Rajpurkar in a newly published perspective in *Nature*, on which he is co-senior author with Eric Topol of the Scripps Research Institute and colleagues from Stanford University, Yale University, and the University of Toronto.

Generalist medical AI is on the cusp of transforming clinical medicine as we know it, but with this opportunity come serious challenges, the authors say.

In the article, the authors discuss the defining features of generalist medical AI, identify various clinical scenarios where these models can be used, and chart the road forward for their design, development, and deployment.

**Features of generalist medical AI**

Key characteristics that render generalist medical AI models superior to conventional models are their adaptability, their versatility, and their ability to apply existing knowledge to new contexts.

For example, a traditional AI model trained to spot brain tumors on a brain MRI will look at a lesion on an image to determine whether it's a tumor. It can provide no information beyond that. By contrast, a generalist model would look at a lesion and determine what type of lesion it is -- a tumor, a cyst, an infection, or something else. It may recommend further testing and, depending on the diagnosis, suggest treatment options.

"Compared with current models, generalist medical AI will be able to perform more sophisticated reasoning and integrate multiple data types, which lets it build a more detailed picture of a patient's case," said study co-first author Oishi Banerjee, a research associate in the Rajpurkar lab, which is already working on designing such models.

According to the authors, generalist models will be able to:

* Adapt easily to new tasks without the need for formal retraining. They will perform the task by simply having it explained to them in plain English or another language.
* Analyze various types of data -- images, medical text, lab results, genetic sequencing, patient histories, or any combination thereof -- and generate a decision. In contrast, conventional AI models are limited to using predefined data types -- text only, image only -- and only in certain combinations.
* Apply medical knowledge to reason through previously unseen tasks and use medically accurate language to explain their reasoning.

**Clinical scenarios for use of generalist medical AI**

The researchers outline many areas in which generalist medical AI models would offer comprehensive solutions.

Some of them are:

* **Radiology reports.**

Generalist medical AI would act as a versatile digital radiology assistant to reduce workload and minimize rote work.

These models could draft radiology reports that describe both abnormalities and relevant normal findings, while also taking into account the patient's history.

These models would also combine text narrative with visualization to highlight areas on an image described by the text.

The models would also be able to compare previous and current findings on a patient's image to illuminate telltale changes suggestive of disease progression.

* **Real-time surgery assistance.**

If an operating team hits a roadblock during a procedure -- such as failure to find a mass in an organ -- the surgeon could ask the model to review the last 15 minutes of the procedure to look for any misses or oversights.

If a surgeon encounters an ultra-rare anatomic feature during surgery, the model could rapidly access all published work on this procedure to offer insight in real time.

* **Decision support at the patient bedside.**

Generalist models would offer alerts and treatment recommendations for hospitalized patients by continuously monitoring their vital signs and other parameters, including the patient's records.

The models would be able to anticipate looming emergencies before they occur. For example, a model might alert the clinical team when a patient is on the brink of going into circulatory shock and immediately suggest steps to avert it.

**Ahead, promise and peril**

Generalist medical AI models have the potential to transform health care, the authors say. They can alleviate clinician burnout, reduce clinical errors, and expedite and improve clinical decision-making.

Yet, these models come with unique challenges. Their strongest features -- extreme versatility and adaptability -- also pose the greatest risks, the researchers caution, because they will require the collection of vast and diverse data.

Some critical pitfalls include:

* **Need for extensive, ongoing training.**

To ensure the models can switch data modalities quickly and adapt in real time depending on the context and type of question asked, they will need to undergo extensive training on diverse data from multiple complementary sources and modalities.

That training would have to be undertaken periodically to keep up with new information.

For instance, in the case of new SARS-CoV-2 variants, a model must be able to quickly retrieve key features on X-ray images of pneumonia caused by an older variant to contrast with lung changes associated with a new variant.

* **Validation.**

Generalist models will be uniquely difficult to validate due to the versatility and complexity of tasks they will be asked to perform.

This means the model needs to be tested on a wide range of cases it might encounter to ensure its proper performance.

What this boils down to, Rajpurkar said, is defining the conditions under which the models perform and the conditions under which they fail.

* **Verification.**

Compared with conventional models, generalist medical AI will handle much more data, more varied types of data, and data of greater complexity.

This will make it that much more difficult for clinicians to determine how accurate a model's decision is.

For instance, a conventional model would look at an imaging study or a whole-slide image when classifying a patient's tumor. A single radiologist or pathologist could verify whether the model was correct.

By comparison, a generalist model could analyze pathology slides, CT scans, and medical literature, among many other variables, to classify and stage the disease and make a treatment recommendation.

Such a complex decision would require verification by a multidisciplinary panel that includes radiologists, pathologists, and oncologists to assess the accuracy of the model.

The researchers note that designers could make this verification process easier by incorporating explanations, such as clickable links to supporting passages in the literature, to allow clinicians to efficiently verify the model's predictions.

Another important feature would be building models that quantify their level of uncertainty.

* **Biases.**

It is no secret that medical AI models can perpetuate biases, which they can acquire during training when exposed to limited datasets obtained from non-diverse populations.

Such risks will be magnified when designing generalist medical AI due to the unprecedented scale and complexity of the datasets needed during their training.

To minimize this risk, generalist medical AI models must be thoroughly validated to ensure that they do not underperform on particular populations, such as minority groups, the researchers recommend.

Additionally, they will need to undergo continuous auditing and regulation after deployment.

"These are serious but not insurmountable hurdles," Rajpurkar said. "Having a clear-eyed understanding of all the challenges early on will help ensure that generalist medical AI delivers on its tremendous promise to change the practice of medicine for the better."

https://www.sciencedaily.com/releases/2023/04/230427114526.htm