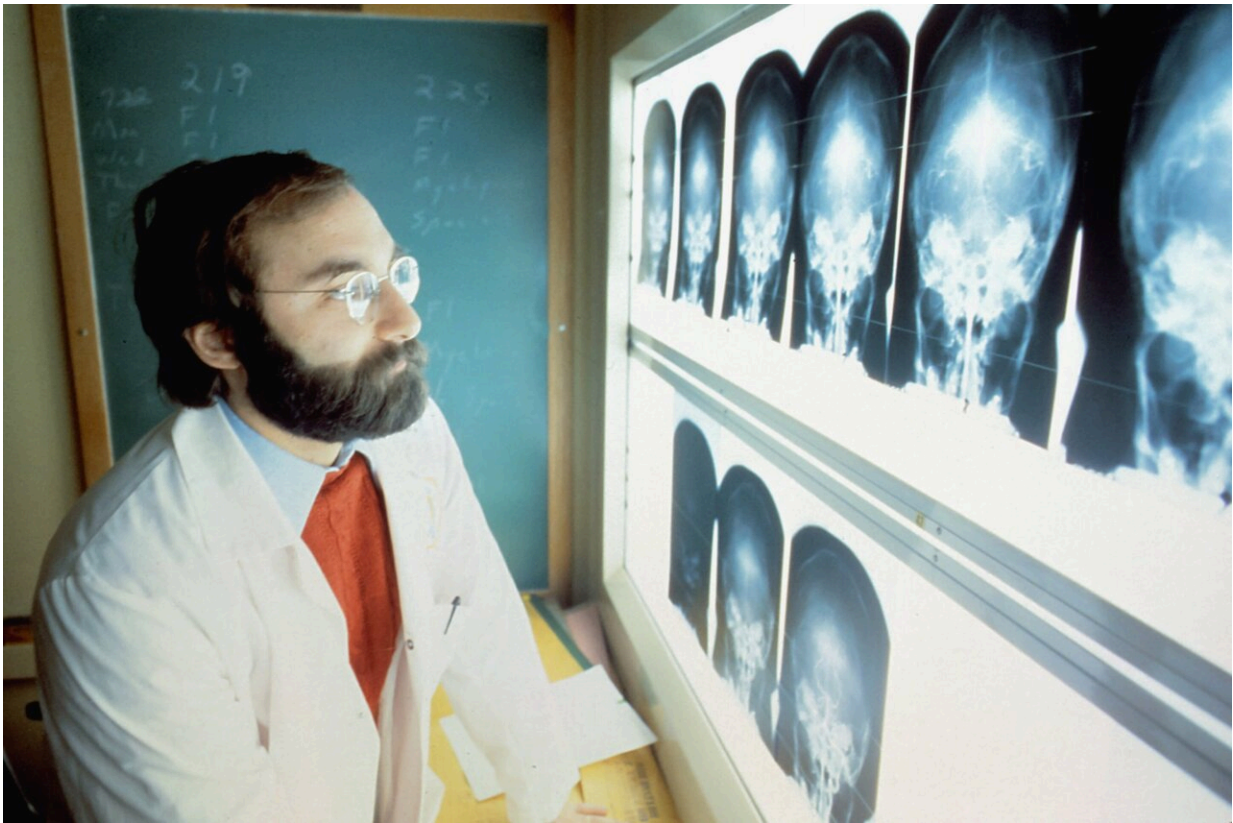


Q&A: The quest to develop fair and ethical algorithms in medical imaging

June 8 2023, by Karen Olsen



Credit: Unsplash/CC0 Public Domain

We're in the midst of an AI boom.

AI is everywhere. If you've used autocorrect when typing or texting,

looked for answers online using a search engine, or used facial recognition software to unlock your phone, then you've been using AI. The technology is ubiquitous and is evolving at breakneck speed.

The use of AI in the health care setting dates back more than 50 years. While chatbots like ChatGPT have garnered recent attention, the first medical chatbot, named ELIZA, was developed in the 1960s. Other AI tools soon followed: In the 1970s, a system dubbed MYCIN was designed to diagnose bacterial infections and recommend appropriate antibiotic treatment, and in the 1990s, the ImageChecker system was developed and became the first FDA-approved image AI system for the computer-aided detection of screening mammograms. Today, medical AI continues to advance, with basic science abilities like predicting the 3D shape of any protein to clinical applications like assisting with surgical procedures.

One of the first medical fields to deploy AI broadly was radiology, where machine learning algorithms are used to help interpret medical images. Medical images contain enormous amounts of data. These data can be mined for patterns and associations that can facilitate disease detection, predict clinical outcomes, and drive therapeutic decisions. But the development of algorithms to assist with these tasks has yet to be fully streamlined, and many questions remain: How are the algorithms regulated? Are the data secure? Are the models developed without bias?

We recently spoke with Maryellen Giger, Ph.D., a professor of radiology at the University of Chicago and the lead contact researcher of the Medical Imaging and Data Resource Center (MIDRC), a data repository that was developed during the COVID-19 pandemic. She told us about the creation of MIDRC, how the imaging repository and resources can be used to develop and test [medical imaging](#) algorithms, ways that bias can be introduced—and potentially mitigated—in medical imaging models, and what the future may hold.

What prompted the creation of MIDRC?

The use of big data has exploded in recent years. But the power of big data can only be fully realized once the data itself has been checked and curated. As the field stands now, medical imaging data used to build clinical algorithms can lack the quality, quantity, and diversity needed to develop ethical and trustworthy products. The downstream effects of such artificial intelligence systems could have enormous ramifications, subjecting underrepresented populations to potential harm and perpetuating disparities in the medical setting.

At the end of 2019, NIBIB and its collaborators were searching for a use case—a specific clinical problem, like how to improve breast cancer diagnoses using mammogram images, for example—to spearhead the development of a robust medical imaging repository. This data repository would need to have diverse and curated imaging data paired with patient outcomes that could be freely accessed by researchers across the globe to facilitate the development of trustworthy machine learning algorithms.

And then the pandemic hit. Seemingly overnight, we had a use case—an urgent need to collect medical images and use AI to predict COVID-19 diagnosis, severity, and outcomes. To do that, we needed to create the infrastructure to collect, harmonize, and store huge amounts of images as well as associated clinical and demographic data. And so in August 2020, MIDRC was born.

At the time of this writing, we've collected chest images (both X-rays and CT scans) from nearly 55,000 patients across the U.S. These images have helped to create 27 in-house algorithms for the detection, diagnosis, monitoring, and prognosis of COVID-19.

How does MIDRC facilitate algorithm development?

The first step in building a good [algorithm](#) is finding good data. During the pandemic, when people were scrambling to develop algorithms, information was pulled from all sorts of sources, resulting in data sets that contained low-quality images, duplicate cases, or patient groups that were not representative of the general population. Garbage in, garbage out: In machine learning, you can't have a good output if you don't have a good input.

MIDRC is working to address this issue in multiple ways. We collect diagnostic imaging data from all over the country and organize these data so that researchers can easily extract what they need to develop an AI algorithm. Medical imaging algorithms are highly specific: they address a specific clinical question for a specific intended population. An algorithm may claim, for example, that it can predict COVID-19 severity from chest X-rays among Black women. The MIDRC database allows researchers to pull the specific data that they need to build an algorithm for their specific clinical use.

Beyond providing data for algorithm development, MIDRC also facilitates algorithm testing and validation. While approximately 80% of the data that we ingest are freely accessible to the public for research use, we keep the remaining 20% of the data in a separate, sequestered data set. This sequestered data set will never be used for algorithm development, as it will be only used for testing algorithms. Now that MIDRC is in its third year, our sequestered data set is large enough to begin evaluating how well a given algorithm is performing. A researcher can come to MIDRC with an algorithm that they've built, and we can extract the intended population from the sequestered data set and evaluate the algorithm to see if it works as anticipated. We've just begun working with companies to evaluate their algorithms and expedite the process of receiving regulatory approval from the FDA for clinical use.

How can bias be introduced into medical imaging algorithms?

Unfortunately, bias can sneak into an AI model in multiple different ways. In our [recent paper](#), we outline about 30 separate factors that can introduce bias throughout the machine learning pipeline. In short, bias can be introduced in any and all steps of the process—from the data that is used to develop the algorithm to how the model is tested and ultimately deployed in the clinic. An algorithm that was developed with one or more sources of bias could ultimately affect patient care.

Some sources of bias seem obvious. An algorithm that was developed using data collected from white patients may not perform as described in a hospital system where the majority of patients are Hispanic. Models that can predict COVID-19 prognosis from chest X-rays shouldn't be expected to predict prognosis for lung cancer using the same type of scan. And algorithms that use images from the same patients to both develop and test the algorithm can't be expected to be generalizable and perform accurately across multiple patient populations.

Other sources of bias are more subtle. One example is temporal bias, where the data used to develop a [model](#) becomes outdated. Think of the early days of the pandemic—when patients checked into the hospital, they were presenting with acute and severe disease without any prior immunity. Today, if the same patient got COVID-19, they've likely had a vaccine or a previous bout of the disease, and the characteristics on their chest X-rays might look a lot different. Different variants throughout the pandemic have led to different disease presentations. In this way, data acquired in 2020 may no longer be as suitable for COVID-19 algorithm development as they were previously—and algorithms developed several years ago will likely need to be re-evaluated to see if they are still performing as intended.

Fortunately, there are ways to mitigate biases in AI models that have been developed to analyze medical images. MIDRC has created an online bias awareness tool that describes potential sources of bias, how these biases arise, and suggestions for how to best correct them. We hope that researchers will use this tool to better understand how bias can be introduced into their development pipelines and take actions to diminish or eliminate biases in their algorithms.

What is the future of MIDRC?

The overall goal of MIDRC is to support the medical imaging AI ecosystem. We've built the infrastructure to house and organize medical images, we've collected a huge amount of real-world imaging data, and we've put forth a concerted effort to educate users about the development of algorithms and potential sources of bias. The next big push is to implement the sequestered data set to enable algorithms to get through the FDA approval process.

As the pandemic continues, MIDRC now gathers data from patients with long COVID and is extending its interoperability with a number of other data repositories. Beyond collecting X-ray and CT images of the chest, we're starting to collect images of the brain and heart captured by additional imaging modalities, such as ultrasound and MRI. Now that MIDRC is incorporating multiple types of images from multiple different regions of the body, it will be poised to pivot to other diseases and clinical problems beyond COVID-19.

Even though the pandemic was the catalyst to build and launch MIDRC, it's more than a repository of COVID-19 images. The collaborations and infrastructure that have been established provide a solid foundation for the creation of more medical imaging data sets and the development of AI algorithms for all sorts of use cases.

The end of the COVID-19 pandemic doesn't spell the end of MIDRC's potential. With the infrastructure built, this could just be the beginning.

Provided by National Institute of Biomedical Imaging and Bioengineering

Citation: Q&A: The quest to develop fair and ethical algorithms in medical imaging (2023, June 8) retrieved 6 May 2024 from

<https://medicalxpress.com/news/2023-06-qa-quest-fair-ethical-algorithms.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.