As AI use expands, ethics at the leading edge

Anne Paxton

February 2024—Artificial intelligence is sizzling, so much so that *New Yorker* magazine, evoking the dazzling and the potentially devouring nature of AI technology, tagged 2023 as "The Year A.I. Ate the Internet." One respondent to a CAP survey of its members on AI called these "exciting but uncertain times."

Commercially available digital pathology platforms already use AI software, and large language models such as ChatGPT, Bard, and Copilot play a growing role in extending AI beyond direct patient care to training and research.

"But there is a degree of angst and concern among pathologists overall about the utilization of AI, particularly as we move forward into education. And there's even less knowledge about it in research," says Suzanne Zein-Eldin Powell, MD, professor of pathology and genomic medicine and director of the anatomic and clinical pathology residency and neuropathology fellowship at Houston Methodist Hospital. She is chair of the CAP's Ethics and Professionalism Committee.

"It's unclear to what extent we want to encourage our trainees to use AI now," says Neil Anderson, MD, D(ABMM), director of the anatomic and clinical pathology residency program at Washington University School of Medicine in St. Louis and a member of the CAP Ethics and Professionalism Committee.

"Some people are very much against it and don't feel comfortable at all," he continues, "whereas others see AI tools for their potential and want us to be using them in a controlled manner."

Fears that AI could become impossible or difficult to control are justified only if humans do not impose controls when AI systems are built, says Brian R. Jackson, MD, medical director for business development at ARUP Laboratories and a member of the CAP Artificial Intelligence Committee.

"As soon as we start talking about AI as being too big to handle or too risky, then what we are saying is that we need to put more effort into the control side of it and that maybe we've gone too far on the autonomy side," he says. "But if pathologists are in the driver's seat and AI is developed in ways that empower and let pathologists do a better job with their existing work, then I think we'll be on a good trajectory."

Drs. Jackson, Anderson, and Powell led a CAP23 course on AI and ethics and spoke with CAP TODAY recently.

The CAP in August 2023 sent an online survey to a random sample of CAP fellows with five to 30 years in practice, all House of Delegates members, and all members of the CAP Engaged Leadership Network (graduates of the CAP's Engaged Leadership Academy course).

The aim was to assess familiarity with and the use of AI diagnostic tools, AI policies and guidelines in place within the laboratory and hospitals/organizations, and AI in training, and to reveal concerns about or views of AI.

The CAP had 152 responses to its survey (2,043 distributed, 7.4 percent response rate), which revealed the following: Twenty percent of academic and nonacademic hospital-based respondents said AI diagnostic tools are being used in their practices or laboratories. Fifteen percent reported they have validated at least some AI diagnostic tools, with academic hospital-based respondents more likely to have validated AI tools (19 percent) than nonacademic hospital-based respondents (13 percent).

Forty-six percent reported their hospital/organization has a policy or process to identify when informed consent is necessary, but 43 percent were unsure whether guidelines or policies are in place to govern or guide the use of AI. Fifty-nine percent were unsure whether their hospital/organization has a policy for data sharing with commercial AI developers.

Many reported they were unsure whether AI was used in their training programs, though 16 percent reported their trainees are using AI tools for presentations and clinical reports, among other things. Sixteen percent indicated their training program offers education on the appropriate use of AI tools. Very few reported that their hospital/organization provides guidance on how to cite the use of AI in manuscript/data preparation. Few reported there is a policy to govern/guide the use of AI in their training program.

Dr. Anderson

That use of AI in medicine is so new explains some of the survey respondents' uncertainty about policies, says Dr. Anderson, who is associate professor in the Washington University Department of Pathology and Immunology. "One of the important things that came out is that a lot of people answered 'I don't know' to a lot of questions. That says that many people are learning about things like ChatGPT and AI from the media and their friends rather than talking with their colleagues about how it can be used. Everyone should know whether or not they have a policy regarding AI, and if they don't they might consider drafting one." Dr. Anderson says guidance, recommendations, and protocols are needed for validating and verifying AI tools, "from informed experts who understand both the laboratory medicine side and the technology."

AI has great promise as a check on plagiarism, he says, either by trainees or researchers. "Maybe it's okay for AI to generate the outline and you fill in the pieces. One could argue that you're still ultimately responsible for what comes out of the AI models. So it's kind of like using Google or Wikipedia to inform your presentation." On the other hand, he adds, "It's not as clear to the average layperson what is going into these models. And if you're using them to generate patient notes or presentations you will use to teach others, is that material even going to be correct? That comes back to someone needing to vet it. If we bypass that part of our training of our residents, we could have a real issue."

That points to the need for more standardized validation protocols for AI tools. In non-AI–based routine testing, "we have a very specific playbook we have to follow when validating a new test, and I don't necessarily know how well developed that is for AI tools at this point because they're all new and cutting edge," he says.

If the data going into the model are not understood completely, we may not understand if the model is working appropriately, Dr. Anderson says. "And if we're using a test to make these higher-order decisions, you want to have a handle on that. For instance, if a model is based off of many different laboratory values that feed into it, what happens when you change one of those tests and don't consider its impact on the model? With QA and QC in the laboratory, you're making sure a test still works. You need to have those same checks and balances with AI."

He cautions too about the need to prevent bias, saying that certain tools, depending on how they're built, might be susceptible. "If the tool is constantly normalizing your data and always giving you answers based on what is most often correct, you're creating a bias. And when you get something that doesn't necessarily fit your model, then it might be inaccurate. If I tried to design some sort of AI model based on one patient subtype, it may or may not work in patients from a different demographic or a different geography."

Residents and fellows need to have a basic understanding of AI tools that are built on laboratory output data, Dr. Anderson says.

There needs to be regulatory oversight also, by people who understand not only the information science behind it but also the laboratory medicine science and clinical science behind it. "Where we run into trouble is when we have people who may have only one type of expertise and not the other types vetting these tools and seeing

whether they're ready for prime time or not," Dr. Anderson says. He suspects that the technology at this time is outpacing the regulation. "That's not altogether surprising, and the regulation will catch up, but that's where we are right now."

The Food and Drug Administration has a role to play in regulating AI for use in medical care, Dr. Jackson notes—at least in premarket evaluation.

FDA mechanisms for postmarket surveillance and enforcement are weak, he says. He hopes the FDA is developing mechanisms to do a better job overall in evaluating AI that's embedded in medical devices.

Wu, et al., analyzed 130 of the medical AI devices approved by the FDA between January 2015 and December 2020, using summary documents of each approved device (Wu E, et al. *Nat Med.* 2021;27[4]:582–584). Almost all of the AI devices (126) underwent only retrospective studies at their submission, based on the FDA summaries. None of the 54 high-risk devices were evaluated by prospective studies. Of the 130 devices, 93 did not have publicly reported multisite assessment included as part of the evaluation study. Only 17 device studies reported that demographic subgroup performance was considered in their evaluations.

Dr. Jackson calls this "not reassuring" in terms of the FDA validating AI for pathology.

The FDA, he says, has never asserted authority over electronic health records. "That's one area of protection where I'd like to see the FDA continue to develop and evolve rules and oversight."

Dr. Jackson

Ethically, the proprietary nature of AI algorithms in itself can be a problem, Dr. Jackson says. "If you're going to let an algorithm loose on patients, you need to be able to evaluate its accuracy and safety." When companies shield their software as intellectual property and make it hard to evaluate, the consequences can be serious, as Michigan Medicine's external validation cohort study of the Epic Sepsis Model (ESM) revealed (Wong A, et al. *JAMA Intern Med.* 2021;181[8]:1065–1070).

In that study, the ESM was found to have poor discrimination and calibration in predicting the onset of sepsis at the hospitalization level. When used for alerting at a score threshold of six or higher (within Epic's recommended range), it identified only seven percent of 2,552 patients with sepsis who were missed by a clinician (based on timely administration of antibiotics).

"Owing to the ease of integration within the EHR and loose federal regulations," the authors write, "hundreds of US hospitals have begun using these algorithms."

The ESM story is an object lesson in the need to control AI, Dr. Jackson says. "Epic has a long history of making its software difficult to evaluate. They see it as an intellectual property area. And it's hard to critique and evaluate it in the public sphere. We need open evaluation, open monitoring, a lot of transparency, and those mechanisms aren't well defined yet."

Some of the most egregious cases of AI ethics problems are in the medical insurance space, he says, where companies use algorithms to deny care. "Historically, they've hired doctors to make those determinations, but it's cheaper to hire algorithms, and they're usually not designed to answer questions about why they rejected a care decision as medically unnecessary. So they can blame the algorithm, rather than take ownership for basically denying legally required care, which is completely unethical."

Dr. Powell

With transparency, some of these outcomes could be avoided, Dr. Jackson says. “But there’s very little transparency. These algorithms are not being independently evaluated for accuracy or performance. There’s no monitoring going on and patients are getting hurt, because the companies are pushing the envelope to see what they can get away with.” It would be much more ethically defensible if the algorithms were implemented to make it transparent to the stakeholders what’s going on, he says.

“As human beings and as organizations and companies, we need to own our responsibilities for developing and implementing and using AI in ethical ways. It’s not okay, if something goes wrong, to point the finger at the AI and say, ‘Oops, the algorithm screwed up.’ No—someone used that algorithm and someone needs to be held accountable if they didn’t put the controls in place to make sure it was going to be used effectively and safely.”

AI model cards, which explain a machine learning model for the purpose of transparency and accountability in development and use of AI, are one way to bring control. “The idea is that you document the algorithm, with some explanation of how it performs, so it adds a level of transparency. It’s not sufficient; it’s a step in the right direction,” Dr. Jackson explains.

He uses car safety as an analogy to the needed protections in software. “Brakes don’t solve the problem alone; neither do airbags or driver’s licenses. But you put them all together, it makes a pretty effective safety network. So the model card is one proposal that would ensure a bit more transparency in how the models perform.”

The classic principles of medical ethics, named in the Belmont Report for the protection of human subjects of biomedical and behavioral research, are patient autonomy, beneficence, nonmaleficence, and justice. “All of those relate to AI in one way or another,” Dr. Anderson says.

Patient autonomy, for example, should affect how patients’ clinical data goes into the AI models, he explains. “Is your clinical data being used in a way that’s right and fair to the person? Beneficence involves making sure results are correct.” Nonmaleficence, the principle of “Do no harm,” could relate to making sure there are checks and balances against commercial interests of corporations overriding ethical treatment, he says.

As for justice, Dr. Anderson sees accessibility as a critical component. “As AI tools become available, to whom will they be available? If we can only have them in areas or hospitals that are relatively well funded with a lot of the expertise, what happens to the hospitals that don’t necessarily have five informatics experts? How do we make sure everyone can take advantage of technologies that have clear benefits?”

The White House’s Blueprint for an AI Bill of Rights and the AMA’s guidelines to advance AI in medical education through ethics, evidence, and equity point to some of the ethics glitches or gaps having drawn attention, Dr. Jackson says.

At the CAP, too, members of the AI Committee have developed a document outlining ethics principles, which members of the Ethics and Professionalism Committee are vetting now. “But it’s very early stage,” Dr. Jackson says.

Dr. Powell cites a study of 253 articles on AI ethics in health care whose authors propose a responsible AI framework that encompasses five main themes for AI developers, health care professionals, and policymakers (Siala H, et al. *Soc Sci Med.* 2022;296:114782). Summarized by the acronym SHIFT, the themes (and some of the

subthemes) are as follows:

- Sustainability: responsible local leadership; societal impact on well-being of humans and the environment.
- Human-centeredness: embedding humanness (recognition and empathy, for example) in AI agents to meet ethics of care requirements; the role of health professionals to maintain public trust.
- Inclusivity: inclusive communication (patient-provider) and involvement in AI governance.
- Fairness: alleviating algorithmic and data bias; health disparities in low-resource settings.
- Transparency: safeguarding privacy; explainable AI-driven models and decisions; informed consent for data use.

In Dr. Powell's view, pathologists are uniquely suited to helping medicine skirt some of the dangers of AI. "From a pathologist's standpoint, of course, we would all be horrified that any sort of algorithm would be utilized clinically without having been rigorously validated at one's own institution. We just need to remind everybody of that. It's the basis of all clinical testing."

Dr. Anderson is especially interested in the risks and potential benefits of AI in training, and he points to a few of the questions it raises:

- Are trainees using AI in a HIPAA-compliant way?
- Are trainees using AI in a way that is consistent with the educational mission?
- Are they using it in a way that avoids plagiarism, and in a way that is safe for patients?

When Google searches were rolled out decades ago, he recalls, the initial reaction from some quarters was "to tell people not to Google stuff, that that was a really awful way and you should consult a textbook if you want the right answer. And now that's silly."

Large language models like ChatGPT could well follow the same course in terms of the degree to which they'll be adopted, he says.

"They will find their way into our day-to-day work in some way, shape, or form. I have no doubt. It's just a matter of to what extent and how they're being used."

Anne Paxton is a writer and attorney in Seattle.