# How Pathologists Can Leverage AI to Improve Patient Care

October 20, 2023

**Lisa Tomcko:**

Welcome to the latest edition of the College American Pathologists CAPcast. I'm Lisa Tomko, content specialist with the CAP. Artificial Intelligence, also known as AI, is streamlining processes within health care, particularly related to diagnosing and managing patient care. In this interview with Becker's Healthcare, CAP member and pathologist M.E. de Baca, who goes by Doc for short, discusses the complexities of integrating AI into patient care, giving consideration to the practical, ethical and collaborative aspects that need to be addressed for effective implementation and improved care. So with no further ado, I'll let them take it away.

**Erika Spicer Mason:**

Hi everyone. This is Erika Spicer Mason with Becker's Healthcare. Thank you for tuning in to today's featured session, where we'll discuss how to harness technology to improve diagnostic medicine and how pathologists can leverage artificial intelligence to improve patient care. I'm thrilled to be joined by Dr. de Baca, who's the chair of the Council on Informatics and Pathology Innovation at the College of American Pathologists.

Dr. de Baca welcome. Thank you so much for joining us today. I really appreciate it.

**Dr. de Baca:**

Oh, thank you, Erika, for hosting this podcast. This is a topic that's very near and dear to my heart. Not only do I chair the CAP Council on Informatics and Pathology Innovation, which I might refer to during the podcast at some point as CIPI, but I also sit on the Board of Governors for the CAP. I'm an anatomic and clinical pathologist with specialty certification in pathology, and I have over 20 years of practical experience in information.

I'm really interested in interoperability and standards in general, but especially reporting standards and in emerging technologies. I've been on the CLIA Advisory Committee, and I'm a past president of the Association for Pathology Informatics. In my day job, I am the vice president for medical affairs at Sysmex America, and I continue to sign out cases with specific pathology partners in Seattle where my wife and I live.

**Erika Spicer Mason:**

Amazing. Thank you so much for giving our listeners a little context and sharing some more about yourself. Sounds like you are the perfect person to talk to us today about the topic at hand technology and AI And I'm really interested to hear about how this applies to the field of pathology. You know, it really feels like AI has kind of entered every nook and cranny of the health care industry.

So looking forward to getting your insights here. So to get us started, I'm wondering if you can share with us what you see as AI's potential for improving diagnostic accuracy and efficiency and just taking that a bit further? In what ways can AI powered image analysis and pattern recognition algorithm really help pathologists in both detecting and diagnosing rare and complex diseases, which then would hopefully enable early intervention and also personalized treatment?

**Dr. de Baca:**

Well, thanks, Eric. And that's a that's a question and that offers an awful lot of options or numbers of paths to respond to. So maybe we should start by agreeing on some definitions. There might be listeners who saw the word AI but don't really have a really good basis or history in studying this. And it would be confusing if I start throwing terms around that I haven't defined.

So the term artificial intelligence is actually kind of a super ordinate for several different models that all have the same kind of goals. So if we look at the ones most used in medicine, machine learning describes an AI method where the system is provided input data and calculates or learns how to process the data to solve the problem, provide it, and machine learning can be divided into three different groups depending on how we try to tell the computer what we want it to do.

So there's supervised learning, unsupervised learning and reinforcement learning. Supervised machine learning is generally used to classify data or to make predictions, and it uses labeled data. So we train the computer by telling it what's incoming, and then later on we throw it unknowns and say, Is this what you used to what you've seen before, or is this something different?

And this is what's used in image analysis and it's widespread in radiology and pathology, unsupervised learning. Just so your listeners know, is used to understand relationships without data set. So within data sets and it understands patterns and it predicts output, but you don't use label data to train the computer. So this is helpful to search for unknown similarities and differences in data to create groups.

And this is widely used, for instance, categorizing users based on their social media activities and reinforcement learning is neither supervised nor unsupervised, and it doesn't require label data or training sets. So think robotics. But I could wander way out of your question, so let me rein myself back in. Let's go to some examples. I said we have supervised and unsupervised learning and that supervised models are provided with labeled data.

So the models, given the answer, say, I have a whole series of images with pictures of a certain type of cancer, and I put in the image and I put in in another column, I say, this is Cancer X, and I show the computer 100 pictures of cancer X, and I train it on that. And then later on I will hand it a picture of something else.

And I would say, What is this? And if it's cancer X, I hope it would tell me that. And if it's not, it should say this is not cancer X. So this is what we do with image analysis in pathology quite frequently. I realized when I was saying this that I used words like models and algorithms. So I'm going to define those and we'll figure that out.

So algorithms are instructions to be followed in calculations or other operations. And this is kind of the program that tells the computer how to learn to operate on its own and how to process these data. And an algorithm in machine learning is the procedure to run on the data to create the model. And these models are the output of the machine learning algorithm.

It's the product of the task specific data that was processed. So this a model that represents what was learned by the machine learning algorithm. And if I give an example, a decision tree algorithm is a whole series of if and statements or if/then statements and the algorithm results in a model comprising the values from those if/then statements.

Now, you asked how I could assist pathologist and maybe I'll get now to your question. I know there are many opportunities for AI and machine learning to be introduced into the practice of pathology. You ask specifically how an AI powered image analysis and pattern recognition algorithms could assist pathologists in detecting and diagnosing rare and complex diseases. And I'd like to jump up again about a thousand feet and say, you know, AI ML isn't just for imaging or for anatomic pathology.

There are lots of opportunities for AI to be used in improving logistics and create increasing efficiencies or managing assets in laboratories, and those should not be underestimated. AI could be used to assist in test ordering to ensure the application of best practices to reduce test duplications or suggesting ordering enhancements. They could be used to better understand our patient populations to demonstrate unexpected emerging result trends or flag low level instrumentation shifts that might go an otherwise unnoticed.

And an even more exciting is the possibility to use clinical test data to create patient specific prediction models like will this particular patient go from normal to high risk in their diabetes journey? Or will this hospitalized patients need an ICU bed in the next 5 hours, 8 hours, 10 hours. So with this, I think that sort of exemplifies that.

AI ML analysis could be used both in clinical and in anatomic pathology to increase efficiencies. If we look at the anatomic pathology world, image analysis could create work lists that are ordered not by their accession number, but the probability that a case is abnormal. So if my caseload comes in in the morning either let's say serum protein, electrophoresis or prostate biopsy cases, things that used to be in my stack of trays now are in my inbox.

And if these could be ordered by and now I'm using air quotes most likely to need brain power so that the cases that require more work, whether that's mental work or ancillary testing, would be given to me first thing in the morning. And so if these quote most likely to need more brainpower, which by which I mean abnormal cases come to me early in the day, that means I'm not too tired and could probably and would the likelihood of error on my part would be lower.

But it also would allow me to initiate ancillary testing before today's cutoff time, and it could thereby potentially reduce my overall case turnaround time. I could go into lots and lots of examples. I'll give you one more. In surgical pathology, image manipulation could align tissue fragments and needle biopsies often these are just little bits and pieces of tissue that get sort of randomly put on the slide, put in the cassette and so they are they are dispersed randomly.

And when the slide is cut, that's what we see on, on our tissue slide currently with microscopy. But there's no reason with an AI or an image analysis that we couldn't just take those little bits and pieces and artificially project them as if they were linear. And so instead of having to go up and down and all around to look for those pieces of tissue, now all I would have to do is scroll to the right or scroll to the left and this would be helpful in understanding how much tumor volume I have per linear measurement, whether that's in pixels or in millimeters, but that would allow for reduced time in looking for the tissue from fragments and increasing the efficiency of the quantification of the tumor.

So those are some relatively easy ideas that could improve a day in the life of pathologist. There are AI and ML algorithms that may help it with time-intensive, repetitive, and/or time-consuming activities like counting mitosis or finding mycobacteria, and machines are much better at these tasks than humans.

I mean, their minds don't wander and they don't get bored and their stomach doesn't rumble. And so those are just some things that could have a huge impact on pathology of the day in the life of a pathologist. Now, I realize that these aren't sexy, but I think, you know, if you would like me to move into the things that would be like earth shattering opportunities, you know, if you take huge image data sets, these will likely lead to the discovery of things that are outside the relatively narrow capabilities of the individual human brain.

Humans could be pointed in directions that we haven't explored yet. So these data sets could to your point, include all known cases of and then error codes. Again, insert rare disease name here and while finding the data sets would be one of the big first hurdles creating and querying such sets for, you know, morphology, mitotic patterns, presence, or absence of changes to the interstitial matrix, etc. could lead to diagnostic and paradox stick or therapeutic inroads.

And with that, I am going to stop because that was probably the longest first question answer in the history of podcasts.

**Erika Spicer Mason:**

Well, you've made history today, so no, but really this is fascinating, and I do appreciate how you outlined all of those examples because you've really taken some complex topics and concepts and you've boiled them down into ways that are easy to understand. At least I was following. So I know that means our listener as well too. So thank you again for sharing all of that.

And as you were explaining the ways that I can really assist pathologists in their day to day, I was kind of picturing that that person on your team, a human who's always anticipating your needs before you even say it. So I can just see the possibilities that open up when you have a tool like that, right?

**Dr. de Baca:**

I mean, this is like artificial intelligence as your sous chef.

**Erika Spicer Mason:**

Yes, that's exactly it. That's a great analogy. Wonderful. Well, you know, as exciting as best is, I know that any new groundbreaking technology will, of course, come with a learning curve. And, of course, challenges, too. So can you share with us what you're seeing as the key challenges the pathologists and even other clinicians are facing as they're adopting these AI technologies?

And also how can they overcome these obstacles in order to really maximize the benefits of these tools and to ensure a better patient outcomes?

**Dr. de Baca:**

Well, some of the barriers are about as mundane as some of my prior examples. So creating data sets requires technology capacity for sure, and we have some of what we'll need. I'm sure that the tech technical capacities will need to increase, but a priority, we need a new type of communication and new methods of collaboration within the diagnostic world.

There's going to need to be a constant cognizance about the need to share images, and we need to have mechanisms built into the workflow that will allow this to happen. Then there's a need of understanding the limitations of AI technology. I mean, it's important to understand what an AI algorithm can do and what's outside the scope of an algorithm.

So what does a positive result mean? For instance, if an algorithm is designed to be used in neuropathology frozen sections and it gives the results of negative, what does negative mean? Does it mean it's negative for tumor? Does it mean that it's negative for primary brain tumors? Does it mean that it's negative for a subset, a specific subset of primary brain tumors?

Or does that also mean that it's negative for metastatic cancers or and or does that also mean that liver proliferative disorders are ruled out? So like any other test and one thing that pathologists are really kind of good at is understanding exactly what the tool is for. You need to know what things do well, and you should know when something should never be asked to do something.

It's crucial to know and I'm going to just repeat this again. It's crucial to know what questions are being asked and answered and what the limitations are with the implication and implementations of new technologies. We also have new ethical questions with AI. Those, you know, is there bias in the test? We know that in the clinical laboratory we've had calculations for eGFR or that were different for Caucasian and African-American patients, and we've come to realize that those calculations were sort of the incorporation of systemic bias in our laboratories.

And those are errors that we are trying to rectify. But we know that we have examples of bias in our data, and we need to acknowledge that because unless we work diligently to mitigate those issues, then these algorithms are only going to replicate institutional and historical bias. And this could amplify this amplify disadvantages that lurk in our data points.

Furthermore, we need to consider that there are biases that are implicit in the data to which we are blind. So if I say, okay, I think eGFR in African-American patients, we knew that was there and it became clear to us that that might be something that was not scientifically based. And so we're dealing with that. But what about things that we don't know exist?

How are we going to find these and the algorithms going to use all the data? If the data are bad, the results will be accordingly skewed. There are all sorts of questions, you know, who's culpable, are responsible for processing the sensitive AI and for the data collection and who is consenting and making sure that things are pseudo anonymized or de-identified.

Who's taking care of the transparency and the storage and the deletion of all the data? And what happens if the algorithm fails? If a third-party solution is installed and it fails, What's the procedure for correction? What's the procedure for updating and upgrading the system? Who pays for that? I mean, if I if I paid for something that's supposed to do X and it doesn't do X, do I have to pay for another thing all over again?

Is it just like, oops, sorry, and who owns these data? What degree of agency does a patient have? And in the utilization or not of an algorithm and one of the things that I think is really important is who gets to choose which algorithms are going to be used. Another question is what are the regulatory barriers? I can go on.

I can seriously talk about these questions for more time than anyone who's listening would believe. So I'm not going to I'm just going to say, you know, we need to make sure that we are creating a culture that advocates for high quality AI implementations and for absolutely exquisite patient safety. And we have to balance that with trying not to discourage vendors from producing such tools.

In order for this to actually work, we have to talk about these problems or potential problems. And in order to solve them, we have to collaborate and create and discuss. But I think most importantly, physicians and data scientists and payers and regulators and patients have to agree from the very beginning that the endpoint of all of these discussions has to be success.

**Erika Spicer Mason:**

Absolutely. It's really powerful. And all of those bigger questions that you raised, questions on accuracy, bias, reliability, ethics, you know that those are questions that are top of mind for, I think, any sector in the health care industry. They're considering these bigger questions. And so it's I'm sure what you're saying is absolutely resonating with our audience. And I'd like to take those, you know, now that you've touched on those concepts of accuracy and bias, especially, I know that that can also raise questions about validation of AI algorithms.

So can you explain why it's so important that these systems are validated? I know you touched on that a little bit, but maybe we can just take that a little bit further.

**Dr. de Baca:**

Sure. So that the lifecycle of AI development does include training and validation and during model training interpretations or data, scientists are trying to optimize the performance metrics that they've predefined. So these are the mathematical represents of the relationships between the data features and the label that we gave that in supervised learning or among the features themselves that we were talking about unsupervised learning.

So once the data are prepared and cleaned up and make, you know, you've made sure that your data are ready for use, then the features are selected and the metrics are chosen. So what? What are we going to show the model and then what do we want it to measure? And once that's been determined, then the model is trained using a training data set.

And there are different ways to define what that training dataset is, and that's outside of the scope of today's conversation. But if anyone is interested, I urge you to go look for that. On Google, there's more information that we will ever believe. So we said once, once the features are selected and the metrics are chosen, the models trained using that training dataset.

And this training is iterative and there's continuous tweaking to the algorithm, basically basing it on the comparison of the prediction model with the labeled data. So if I'm putting N this is cancer X, this is cancer X, this is cancer X, this is cancer X, and then I am looking at it and I don't I can shift things so that the certainty of that being true is to my liking.

And in the final steps of development, then this model with its now optimized settings, is trained using the entire training dataset. So I've held some stuff, I've held some data back and after having gone through this iterative process, then I send in more data that I control. I know that this is, you know, 57 cases of cancer X and 12 cases of something else.

And I now run the new data through the model and I test it to see what the data comes out. And it's similarity in average performance against the training data. So validation is an activity that ensures that the end product stakeholders needs and expectations are being met. And in other words, the validation process is for assessing whether or not the final product is working as it's supposed to, or in this case, quantifying the tumor and saying, you know, whether or not it was doing that correctly.

So this is really important because you need to know what you're validating and you need to know what your sets are. That's the initial process before this would go out into the real world or into the clinical application. So validation data for imaging studies, for example, need to be representative of the target population and you need to consider a whole bunch of things like geographic factors and time factors into disease prevalence or population specificities, racial and gender diversities.

And then there are the basic system things like the camera settings or the specifications of the images or the acquisition specifications.

**Erika Spicer Mason:**

Yeah, Thank you so much for sharing about that. And you know, going back to something you were saying about the part of the process where someone will be running the new data through the model, is that where pathologists kind of have a role in this validation process? Or maybe I have the wrong idea. Could you share a little bit more about that?

**Dr. de Baca:**

Of course. You know, it occurs to me maybe I need to make sure we're using the same terms again. We're seeing validation, and I probably used the word verification. I'm not sure if I did, but there's a validation stage and a verification stage and there similar, but they have very different roles. So in case I did that, I'm going to let people know.

And if I didn't, then this is this is a preemptive strike. So validation, like I said, that that informs about the way that the model, the method is performing and the verification then demonstrates what the local input implementation results are in line with how that test was designed to perform. So because deep learning AI or ML identifies patterns, the data used to train those models need to be annotated and validated by experts.

We call this labeling before annotation and labeling are basically the same thing. So if we were to exclude pathologists during the development and this validation process, we run the risk of creating tools that are based on faulty non-medical people, developer based assumptions. And you know, of course there is additional time and additional cost to having an expert informed AI model perfected, but there is also a non-trivial degree of unawareness of the complexity of many of the diagnostic problems that pathologists face.

If someone who doesn't understand the media is the one who's creating the tool. So I would assert that pathologists need to be involved in the development of tools from the very beginning so that those can be deployed with confidence to assist in pathology practice. Now, once a laboratory has purchased an item or an institution has purchased a tool, that's going to have implications on the laboratory workflows or on laboratory data, then I would also assert that pathologists must be involved in that in the verification process, and that's to ensure that the tool sits properly within the workflow and that all of the variables be that pre, analytic, analytic or post analytic that could have an impact on the algorithm are considered and that the verification to the site specific model implementation compares to the test as it was previously validated.

So pathologists need to play an integral part in both of these steps where we're the only ones that understand the pathology question that is being asked, and we're the only ones who understand the workflows and the patient population characteristics in our own laboratories.

**Erika Spicer Mason:**

We got it. So pathologists basically need to be involved in both the validation and verification. They're crucial to those steps. Well, thank you again for elaborating there. And I know we have limited time today, so I want to keep going down this road, but I feel that we need to shift just so we can cover another aspect of how I can bring benefits to health care organizations.

And so I'd be remiss if we didn't talk a little bit about cost savings. As we know, so many organizations are up against financial challenges right now. So in your view, where do you see I applied in pathology contributing to cost savings and how do you see I reducing, for example, laboratory costs without compromising the quality of care or patient safety, which I know you emphasized earlier?

**Dr. de Baca:**

Yeah. Well, what what I do know and what I think we all know is that change is upon us. And I think that we can see you can assume that some of this technology will help us work more efficiently by actively participating in implementations and validations and verifications. We are going to be assuring quality and trying to minimize risk to the best of our abilities.

What we don't know yet is to what degree these tools are going to populate our laboratories or and change or replace our current systems. And we do not know what the acquisition impact is going to be on our budgets. What I can say is that human agency is needed for technology to work. You know, technology doesn't just cause change by its own or on its own human decisions about implementations or not implementation, not implementing technology lead to what we typically consider to be the impact of technology.

And those choices are really dependent on context. So it's hardly visionary to assert that AI is going to change our practice. But we do need to keep in mind that medical AI provides no explanatory power, and it's it can't search for causes of what's observed, and it might recognize and classify a lesion, but it doesn't infer causation or prevention or treatment or relationships to other diseases.

And a pathologic diagnosis is a considered a cognitive interpretation based on the training, experiences, biases and your mistakes that the pathologists and as pathologists, our professional value arises from our ability to render a clinically relevant interpretation that is influenced not only by hard data, but also the soft, clinical and patient specific information that we assimilate from different human language and communication-based interactions.

So you're asking for ideas about impacts that I don't think I can actually quantify. Currently, there are just so many roads that we could follow and there are so many paths that will lead us into different directions. But I think that AI models are going to take us in directions that are so far inconceivable. But I think that pathologists will still be guiding this pathology starship, and at the end of the day, the patient will be best served knowing it's the pathologist who's now empowered by new AI tools who are going to render these diagnoses.

**Erika Spicer Mason:**

I really appreciate that answer, Dr. de Baca because there is a lot that's unknown still, and I think that you've provided us with some really great insights and what you think is to come. But ultimately there's a lot that's still unknown. So I appreciate you leaving space for that part.

**Dr. de Baca:**

Thank you. One final comment, I guess I didn't mention at the beginning that I trained first in ophthalmology and I lived for a long time in Germany, and so that gave me an appreciation not only for medicine because I love it, but also sort of the German medical history and specifically Rudolf Virchow, who is is sort of the father of pathology.

He's the guy who was an internist and from his understanding of basic science, created out of nothing the specialty of pathology. He created methods to demonstrate from a tissue basis that this was that, that we had a reproducible language, a visual commonality that would help us to diagnose common disease processes that had common features. And one person started a specialty that revolutionized medicine.

People understood the importance of this information and the power that it brought to helping their patients. And this was all based on one person and his bold understanding of the science and ways that that these new methods could be used to impact patient care pathology often is characterized as being a very conservative specialty. And of course, pathologists have to be extremely careful to make the right diagnosis, extremely careful with the safety of their patients.

It's their lives depend on that. But I think that we actually sit in a pivotal moment. Let's call it a virtual moment right now, where if if we use the information that we have, it can take us in places where, like Virchow's colleagues, they probably thought the guy was nuts when but with the right amount of understanding of the problems and with the right amount of caution, pathologists have the opportunity opened a window that could seriously change the way that medicine is practiced.

**Erika Spicer Mason:**

That's a fantastic note to end on. Dr. de Baca. It really is incredible to think about all of the possibilities, and I want to thank you again for your time and giving us insight to that today. I feel like I've served not only as a moderator, but a student too. So it's been a real treat for me to talk to you today. Thank you so much again.

**Dr. de Baca:**

Thank you. Erica is great. Great to get the word out about these possibilities that the world is a fascinating place and we live in fascinating times.

**Erika Spicer Mason:**

Absolutely. We will. Thank you. And I can't wait to hear feedback from our listeners on this topic and the session. So thank you again to Dr. de Baca. And we'd also like to thank our sponsor for today's session, the College of American Pathologists. And to our listeners, I hope you have a great day and thank you so much again for tuning in with us on behalf of the CAP.

**Lisa Tomcko:**

A big thank you to Dr. de Baca and Becker’s Healthcare for this episode on leveraging AI to improve patient care. Stay tuned for future episodes of CAPcast, and for more information about the CAP visit us at cap.org.