

Healthcare AI, Radiology solutions

Evaluating AI-driven analytics for radiology, part 2: What “good” looks like

[Sheela Agarwal, MD MBA](#) | Chief Medical Information Officer, Healthcare Diagnostics

December 2, 2021



A 99.9% accuracy rate looks great on paper. (Or a web page.) But it could be concealing hundreds of time-wasting false positives. Or a biased algorithm that fails to deliver when fed with your own organization's data. In part two of my guide to AI-driven analytics, I explore the performance metrics you're likely to encounter when evaluating different solutions, and the questions you should ask your vendor as you explore their capabilities.

Right now, AI is transforming radiology—whether it's supporting diagnosis and care planning, enabling clinical research, or powering quality improvement initiatives. But the field of AI-driven analytics is often new to radiologists and their organizations.

[In my last post, we explained the technical terms](#) that anyone shopping for a clinical analytics solution must understand to make an informed decision and a smart investment. This time, we'll explore the performance claims you might see, the additional context you should always ask for, and what success has looked like for one trailblazing provider and its patients.

Accuracy demystified

Accuracy tells you how often a model is correct. But that's not very meaningful without some additional context. First, you need to know the sample size that the accuracy rate is based on. Identifying every radiology report with a follow-up recommendation in a sample size of ten is much less impressive than

doing the same with a sample of 10,000.

What’s more, some AI vendors will cast a very wide net when calculating accuracy to capture as many positive results as possible. To continue the example above, if you’re looking to track follow-up recommendations in your radiology reports, a model that includes hedging statements like “follow-up as clinically indicated” could create significant noise—and extra work for your institution.

It’s also important to understand that high accuracy with a particular dataset won’t always translate into real-world performance. A model can be very accurate for one dataset, one institution, or one demographic, but not for another.

Accuracy *is* an important metric—but must be paired with “precision.”

Precision in a nutshell

Precision tells you how many items, out of all the items an AI model has identified, it has identified *correctly*. And again, it’s of limited meaning on its own.

Let’s imagine that you’re still looking for reports with follow-up recommendations, but this time in a sample of 1,000 reports. Your analytics tool identifies ten reports with follow-up recommendations and is right in every case. But it misses the fact that the rest—all remaining 990 of them—had follow-up recommendations too. The AI tool’s precision is 100%, but its accuracy is 1%. If you were to put your faith in *this* analytics tool, you would also be putting your business and patient outcomes at risk.

Recall rate

This is your *true positive* rate—the number of results your AI analytics have identified, divided by the total number of correct results.

Recall rate is going to be invaluable when you’re tuning the performance of your own AI model. However, it’s difficult for major analytics vendors to calculate because it requires a manual review of the dataset to create a ground truth. And if you have a large enough dataset to create an accurate, precise, unbiased solution, manually reviewing all those reports requires a significant investment of time and resources.

Will this analytics solution perform for *you*? Key questions to ask

Before you invest in an AI-driven analytics solution for radiology, you’ll want to know how well it’s likely to fit your people, processes, and needs.

Ask about the sample sizes that any performance claims are based on, and how varied that sample is. Has the solution only been tested on data from hospitals that have very different populations compared to yours, or different ways of reporting on various findings?

You should also ask about how finely you’ll be able to filter results. Will you be able to extract the information your institution needs, and only that information? After all, what you’re interested in will be very different if you’re a radiology group working to clamp down on hedging statements or an academic group looking for over- and under-imaging.

The only way to be sure radiology analytics are going to deliver

Even once you’ve asked a vendor about the accuracy, precision, size, and variety of their datasets, and the filters you can apply, there’s still one question left to ask.

*Can we see how it performs using **our own** data?*

A reputable solutions vendor will say, “Of course!” It’s something that Nuance frequently performs for potential customers. We’ll arrange the necessary agreement, take a sample of your data, evaluate it for you, and compare our AI’s results to your ground truth—so you know *exactly* what you can expect if you proceed to implementation.

What good outcomes look like: In practice

What does a successful analytics implementation look like? Let’s look at how one Nuance customer has built on our powerful analytics capabilities to gain valuable insights and deliver better care.

Summa Health System is a growing healthcare organization in Akron, Ohio that serves more than one million patients every year. [It uses mPower Clinical Analytics to help identify incidentally detected lung nodules, and ensure patients receive the recommended follow-up and care.](#)

Before using mPower’s AI-driven data analytics, an average of eight ED patients per month were referred to Summa’s team of lung navigators with incidental nodules. After deploying mPower, the system was identifying around 60 patients a month—an almost eightfold increase. Crucially, over 30% of these referred cases presented with lung nodules over the 8mm threshold at which the risk of malignancy rises significantly.

Summa Health’s incidental nodule process has not only led to higher follow-up rates but is supplementing its traditional lung screening programs for high-risk populations. They are also expanding the mPower Clinical Analytics capabilities to identify and track dozens of critical results beyond just pulmonary nodules, further contributing to improved patient care across specialties.

Let’s talk AI-driven analytics

We’re happy to answer all your questions about AI-driven analytics for radiology and our industry-leading solutions. Visit [mPower Clinical Analytics for medical imaging](#).

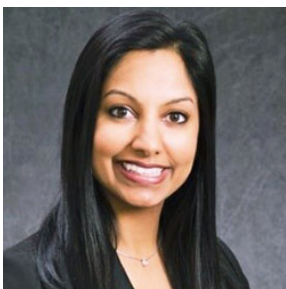
Tags: [Radiology AI](#), [Data analytics](#), [RSNA](#), [Radiology reporting](#), [Natural language understanding \(NLU\)](#)

More Information

Learn more about mPower Clinical Analytics

Explore how our AI-driven analytics solution helps teams monitor, understand, and improve financial and clinical performance.

[Learn more](#)



About Sheela Agarwal, MD MBA

Dr. Sheela Agarwal joins Nuance from Bayer Healthcare Radiology, where she held the position of Digital Medical Advisor, acting as medical lead for the Digital Solutions Business. Prior to that, Dr. Agarwal held other leadership roles at Bayer including Head of the Digital Solutions Business and Head of Medical Affairs for region Americas. She is an experienced leader with a history of working in academics and industry, has written multiple AI-related publications through her work with the American College of Radiology DSI and was recently involved in establishing “AI Central,” the FDA-Cleared Algorithm Catalog. Dr. Agarwal completed her undergraduate and graduate degrees in Economics, as well as her MD from Duke University, her graduate radiology training subspecializing in abdominal radiology and MRI from Massachusetts General Hospital/Harvard Medical School and her MBA from the University of Chicago Booth School of Business. She continues to practice as an abdominal radiologist at Lenox Hill Radiology in New York City.

[View all posts by Sheela Agarwal, MD MBA](#)

