
— DIAGNOSTIC IMAGING LITIGATION

When the Algorithm Saw It First

AI in Radiology and the Missed Finding

BY MICHELLE CARROLL · FOUNDER & CEO, LEXCURA SUMMIT

Radiology is where medical AI has moved fastest. The FDA's public list of authorized AI-enabled devices now runs past a thousand products, and most are radiology tools. Some flag possible brain bleeds or pulmonary emboli on CT. Some find lung nodules or fractures. Some help read mammograms. Some quietly reorder a radiologist's worklist so the scans most likely to be urgent get read first.

Those tools can catch things a tired human misses at 3 a.m. They also change the shape of a malpractice case, and one study suggests they may change how juries see it.

In a 2025 study in *NEJM AI*, researchers at Brown University gave more than 1,300 people hypothetical malpractice scenarios in which a radiologist missed a finding. When participants were told an AI tool had correctly flagged what the radiologist missed, they found for the plaintiff much more often: about 73 percent in a brain bleed scenario, versus about 56 percent when AI wasn't mentioned. But when they were also told how often the AI itself gets things wrong, plaintiff verdicts fell to about a third. How the AI's role gets explained may matter as much as whether it was there.

PIN DOWN WHAT THE TOOL ACTUALLY DID

"The hospital uses AI" tells you almost nothing. Start by identifying the product, the version, what the FDA authorized it to do, how the hospital configured it, and whether it actually ran on this study. Tools get switched off, fail to process images, or are set up to run only on certain protocols.

Then figure out what kind of tool it was, because that shapes the whole case. A triage tool reads the study as soon as it's done and, if it suspects something urgent, pushes it up the list and may send an alert; it doesn't diagnose anything. A detection tool marks where it thinks the abnormality is. A measurement tool sizes a nodule or a bleed. Some tools generate text that ends

up in a draft report. With triage, the question is usually whether the study got prioritized and how fast it was read. With detection, it's whether the radiologist saw the mark and what they did with it. With measurement, it's whether the number was right and whether it changed a recommendation.

A FICTIONAL CASE

A 62-year-old woman gets a chest CT in the ER at 2 a.m. for shortness of breath. The hospital runs an AI tool that looks for pulmonary emboli, and it flags the study as positive. An overnight teleradiology group issues a preliminary report: no PE. She's treated for possible pneumonia and admitted. The next morning the hospital's own radiologist reads the study and finds a clot. By then she has collapsed.

WHAT THE TELERADIOLOGIST SAW

Did the alert even reach the teleradiology group? Many tools are built into the hospital's own system, and a remote reader on a different platform may never see the flag. That's a system failure, not just an individual one.

If the alert did reach the reader, what did it look like: a pop-up, a mark on the image, a line on the worklist? Is there a log showing it was opened? Many systems record whether the AI result was viewed, and that audit data can vanish if nobody asks for it early.

WAS IT THERE TO SEE?

Set the AI aside entirely for a moment. Was the clot visible on the images? How big, and where? Answering that takes a qualified radiologist looking at the images, not the report. The standard for reading a CT doesn't change because a tool was involved. The AI flag may help show the finding was detectable, but it doesn't create the finding.

PRELIMINARY TO FINAL

Overnight teleradiology reads are often preliminary, with a final read by the hospital's radiologist in the morning. Discrepancies happen, and most hospitals have a process for communicating them. How fast was this one caught? Who was told, and when? What did the treating team do?

WHEN THE AI MISSES TOO

Sometimes the tool misses the finding as well. That can help the radiologist, since it suggests the finding wasn't obvious. But it can raise a different problem: automation bias, the well-documented tendency to trust a machine's silence over your own eyes. How long was the study open? Were all the series reviewed? System logs can sometimes answer that.

DOCUMENTATION OF THE RESPONSE

Radiology groups are still figuring out how to document their response to AI flags. A radiologist who disagreed with a flag and wrote down why is in a very different position from one who never acknowledged it. Look for that documentation, and if it's missing, check whether group policy required it.

WHO'S RESPONSIBLE

These cases can involve the teleradiology group and its radiologist, the hospital and its radiologist, the ER physician, the admitting team, and possibly the AI vendor. Claims against developers raise product liability and federal preemption questions that are for counsel and are developing quickly. Clinically, the reviewer can document what the tool was designed to do, what it did here, and whether it performed as expected. The teleradiology contract often spells out turnaround times, preliminary versus final reads, and critical results communication. Hospital AI policies may say whether radiologists must review and document every flag.

CAUSATION STILL HAS TO BE PROVEN

Even if the clot was visible and the flag ignored, you still need to show what an earlier diagnosis would probably have changed. Would anticoagulation have started at 2 a.m.? Given her condition, would something more aggressive have been considered? Would it probably have prevented the collapse? That's for a pulmonologist or intensivist, and it depends on how she was doing overnight: vital signs, oxygen needs, signs of right heart strain, whether she was getting worse. The overnight nursing flowsheets can matter as much as the radiology report.

SCREENING IS ITS OWN ANIMAL

In mammography, AI increasingly works as a second reader. There, the issue usually isn't one dramatic overnight miss. It's whether something the AI marked, or that was visible on a prior study, got acted on over months or years. Those cases look like other delayed diagnosis cases: comparisons with prior studies, recall decisions, follow-up communication, and whether the patient was told about her breast density and what it meant.

RECORDS WORTH REQUESTING

- Images in native DICOM format
- Preliminary and final reports with timestamps
- The AI tool's output for the study and logs of when it was generated and viewed
- Product name, version, and FDA authorization; hospital configuration settings
- Hospital and radiology group AI use policies
- The teleradiology services agreement

- Discrepancy and critical results communication policies, and records of the communication
- Overnight nursing flowsheets and physician notes

SOURCES

- Bernstein MH, et al., study of mock juror perceptions of radiologist liability with AI, *NEJM AI* (June 2025); Brown University news release (July 28, 2025)
- U.S. FDA, Artificial Intelligence-Enabled Medical Devices list

This newsletter is for education only and isn't medical or legal advice. The case is fictional.

ABOUT LEXCURA CLINICAL INTELLIGENCE

Lexcura Clinical Intelligence provides integrated clinical and regulatory analysis for attorneys handling healthcare litigation. If your case involves an AI-flagged or AI-missed finding and you want to know what the tool did, who saw it, and whether the image supported the diagnosis, we'd welcome the conversation.

Michelle Carroll

FOUNDER & CEO, LEXCURA SUMMIT

