

Law360

How Unchecked AI Exposes Expert Opinions To Exclusion

By **Barrington Dyer, Jacob Karim and Curtis Park** (December 3, 2025)

Picture this: Your expert witness used an artificial intelligence tool like ChatGPT, Copilot or Claude.ai to help generate their expert report, resulting in fabricated references, fake credentials or faulty calculations.

Worse, the errors go undiscovered until after service of the report. The cost for this oversight ranges from a mild admonishment from the court to the utter exclusion of your expert's opinions.

And in intellectual property litigation — where reliance on expert testimony to prove liability and damages is almost always a necessity — the exclusion of an expert's opinions can spell the end of the case.

Could this mishap have been avoided?

Much has been made of attorneys' misreliance on AI tools in drafting pleadings, with a Nov. 10 court opinion in *Shelton v. Parkland Health* by the U.S. District Court for the Northern District of Texas observing that "[t]he number of judicial decisions decrying the problems associated with legal briefs prepared using [Generative AI] is substantial and growing weekly." [1]

Less attention has been paid to some expert witnesses' misreliance on AI tools in the drafting of their reports. However, a growing number of cases illustrate the potential for misuse of AI tools by experts in litigation, resulting in reports with hallucinated information or unexplainable analysis. [2]

To embrace the efficiencies AI tools introduce without falling victim to the risks they pose, attorneys and experts should be wary of these common pitfalls and implement a few key best practices.

Lesson 1: The Greater the Error, the Greater the Penalty

A common misstep is relying on AI tools to produce accurate information. In other words, failing to fully appreciate the risk of hallucinations. [3]

In an expert report, these hallucinations may take the form of invented facts, quotes or case names that do not exist. As more experts embrace AI assistance in generating their reports, the risk of hallucinations rises. [4] And as a few real-life examples reveal, the greater the oversight, the greater the penalty.

In *Kohls v. Ellison*, for instance, an expert's use of GPT-4o to draft his declaration resulted in the inclusion of two nonexistent academic citations and an incorrect attribution of authorship. [5] Upon discovering the errors by his expert, the defendant moved to amend his declaration.

While the U.S. District Court for the District of Minnesota credited the defendant's initiative



Barrington Dyer



Jacob Karim



Curtis Park

to correct the errors, in its Jan. 10 order, it underscored the importance of "validating the truth and legal reasonableness of the papers filed" under Rule 11 of the Federal Rules of Civil Procedure, along with the significance of signing a declaration under penalty of perjury.

Stressing that the purpose of the rules is to enable courts to trust the "indicia of truthfulness," the Minnesota court took a stern stance and excluded the expert's testimony entirely, stating in no uncertain terms that it will not "make allowances for a [party] who cites to fake, nonexistent, misleading authorities." [6] In the court's view, the "attorneys and experts [had] abdicate[d] their independent judgment and critical thinking skills in favor of ready-made, AI-generated answers." [7]

Even if an expert's opinion survives exclusion, some credibility may still be lost, as one expert for AI-chatbot-maker Anthropic learned in using the company's own Claude.ai tool to format the citations in her declaration. [8] The tool correctly formatted some citations, but falsified one entirely — a clear example of an AI hallucination.

Upon uncovering the error, the plaintiffs in *Concord Music Group Inc. v. Anthropic PBC* urged the U.S. District Court for the Northern District of California to strike the expert's declaration. Anthropic, in response, argued that the mistake had been overlooked in a manual citation check, but the court was unconvinced.

The court questioned how such an error could have escaped correction from a manual cite check and reiterated that its standing order required counsel to personally verify the accuracy of the contents of submissions. Unlike the fully fabricated sources in *Kohls*, though, Anthropic's fictitious citation linked to an actual article. As a result, instead of full exclusion, on May 23, the court struck only the paragraph containing the hallucinated source, but nonetheless emphasized that the incident "undermine[d] the overall credibility" of the expert's declaration. [9]

When an expert's report is rife with hallucinations, however, the consequences can be devastating, as was the case in *Khoury v. Intermountain Health Care Inc.*, in the U.S. District Court for the District of Utah. [10] There, an expert for the plaintiff alleging False Claims Act violations against an anesthesia practice was asked to opine on the lawfulness of certain charges billed to Medicare and Medicaid programs.

Unbeknownst to the plaintiff, the expert had relied on ChatGPT to generate his report and failed to verify its contents. Not until his deposition did it come to light that he had relied on ChatGPT to help draft his report, which was apparently riddled with AI-generated misquotes and miscitations.

To salvage the suit, the plaintiffs immediately moved to substitute their own expert. The defendants, on the other hand, moved to exclude the expert, sought sanctions in the form of costs and fees, and asked that the litigation be stayed, arguing the "fraudulent report ha[d] thrown expert discovery in this case into disarray." [11]

In August, the District of Utah granted the stay, and before the motions to substitute, sanction or exclude could be decided, the case was dismissed on Sept. 30 without objection from either side. [12]

Altogether, the message is clear: AI tools can add efficiencies, but often at the expense of accuracy. Assistive tasks like drafting outlines, organizing exhibits, locating information, generating tables, proofreading and formatting citations add opportunities for the inclusion of misinformation absent human review.

Lesson 2: Blind Reliance Can Keep an Expert From Opining

Even if an expert's report is free of hallucinations, it may contain analysis performed by AI that the expert cannot sufficiently explain — thereby abdicating the expert's role to AI. Blind reliance on an AI tool without an understanding of how the tool reaches its output — either the analysis it applied or the sources it extracted — can lead to the exclusion of unsubstantiated opinions.

Rule 702 of the Federal Rules of Evidence, for example, allows an expert witness to testify as to their opinion, provided that the testimony is a product of reliable principles and methods.[13] In federal court, where the Daubert standard applies and judges serve as gatekeepers over the reliability of an expert opinion, an expert who lacks an understanding of how the AI output was generated runs the risk of exclusion.[14]

And even in the few state courts that still follow the Frye standard and look to whether experts' methods are generally accepted by the scientific community,[15] undiscerning reliance on AI can render an expert opinion unfit. Two cases exemplify this risk.

In *Jackson v. Nuvasive Inc.*, in the U.S. District Court for the District of Delaware, the defendant's damages expert relied on third-party tools to calculate damages on eight patents.[16] One of the tools, Derwent, calculated "a 'Combined Patent Impact' score which represents the importance of a patent relative to others,"[17] while a second tool, IPLytics (billed as an AI-powered IP analytics tool), calculated a patent's competitive impact.[18]

Use of both of these tools by the expert was found to be unreliable in March by the District of Delaware. The expert's use of Derwent was unreliable because he was unable to explain how the tool calculated the combined patent impact scores.

The court went on to write that the expert's "ignorance of Derwent's algorithm is especially concerning given that Derwent's 'machine learning processes' have seemingly generated an incorrect assignee for the asserted patents via its 'Optimized Assignee' output." [19]

Equally flawed was the expert's use of the IPLytics tool. Even though the expert had demonstrated an understanding of how IPLytics calculated a patent's competitive impact, he simply assigned equal values to IPLytics' input factors without any rationale connecting the factors to a patent's value.[20] For example, a factor such as the number of inventors listed on a patent was, in the court's view, a "dubious" indicator of a patent's value.[21]

In the end, the lack of reliability associated with the expert's opinions led the court to grant the plaintiff's Daubert motion and exclude the expert's testimony.

The Matter of Weber illustrates a different, yet related, reliability issue: Not only can AI-performed analysis be opaque, but the data it draws from may be unknown, and the output it generates inconsistent. Add to that, if the prompts used with the AI tool are not recorded, the problem is compounded.

In *In re: Weber*, the expert for a party challenging the fiduciary administration of a trust relied upon Microsoft Copilot to cross-check his calculation of damages, but could not recall what input or prompt he used to check his calculations.[22] Nor could he state what sources Copilot relied upon, much less explain how Copilot works or how it arrives at a given output.

In its own experimentation with Copilot, the New York Surrogate's Court, Saratoga County,

noted on Oct. 10, 2024, that even the same prompt returned a different value each time. And when presented with the question "are your calculations reliable enough for use in court," Copilot responded, "[w]hen it comes to legal matters, any calculations or data need to meet strict standards. I can provide accurate info, but it should always be verified by experts and accompanied by professional evaluations before being used in court." [23]

As such, the court found "the record ... devoid of any evidence as to the reliability of Microsoft Copilot in general, let alone as ... applied here," [24] and unsurprisingly, the expert's damages calculations were determined to be untrustworthy.

Jackson and Weber show that experts should not overly depend on AI calculations without a firm understanding of how the calculations were reached. Considerations for assessing the admissibility of AI-assisted expert opinions include the expert's comprehension of a tool's algorithm, reliability and acceptability in the field.

In addition, experts who can neither recall the prompts they used nor the inputs they provided to reach their AI-assisted calculations are likely to find themselves unable to defend their opinions.

Actionable Steps to Avoid Risks Associated With Experts' AI Use

To avoid the mishaps that can result from an expert's overreliance on AI tools in crafting a report, consider the following four practices.

1. Discuss the expert's use of AI tools in advance. Cover counsel's expectations of the expert, as well as any policies foreclosing the use of AI in generating a report. Also, ensure that the expert is aware of the risks of using unsecure AI tools on confidential information and is taking appropriate safeguards.
2. Disclose the use of AI as required by the judge or jurisdiction. Many courts, both federal and state, have issued standing orders on the use of AI in court filings — ranging from those simply reminding attorneys of their obligations to verify the contents of any submission, to others requiring the disclosure of AI use in any exhibits or pleadings. A minority of courts, including some in the Northern and Southern Districts of Ohio, prohibit the use of AI altogether.
3. Document any AI tools, prompts and outputs considered in forming expert opinions. To the extent an expert considers AI-generated facts, data or content in forming opinions, that expert should maintain records of the AI tools used, the prompts entered and the outputs extracted.
4. Direct the review of any AI-assisted drafting for accuracy. Last of all, meticulously verify the contents of any expert submission drafted or formatted with the assistance of AI. Treat every AI-assisted sentence and citation as unverified, and check each factual assertion, quotation, citation, calculation and exhibit reference against the ground truth.

Following these simple practices can save your expert's opinions from exclusion and, along with those opinions, save your case.

Willkie Farr associate Rupali Singhal contributed to this article.

The opinions expressed are those of the author(s) and do not necessarily reflect the views of their employer, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.

[1] Order on Rule 11, Shelton v. Parkland Health, No. 3:24-cv-2190, (N.D. Tex. Nov. 10, 2025), Dkt. No. 59 at 6.

[2] See Cade Metz & Karen Weise, A.I. Is Getting More Powerful, but Its Hallucinations Are Getting Worse, N.Y. Times (May 6, 2025), <https://www.nytimes.com/2025/05/05/technology/ai-hallucinations-chatgpt-google.html>.

[3] For the uninitiated, an AI "hallucination" occurs when an AI tool outputs content that sounds plausible but is untrue, fabricated, or unsupported by its sources. When AI Gets It Wrong: Addressing AI Hallucinations and Bias, MIT Management (last visited Nov. 5, 2025), <https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/>.

[4] Metz, *supra* note 1.

[5] Order Granting in Part and Den. in Part Pls.' Mot. to Exclude and Den. Def.'s Mot. for Leave to File Amend. Expert Decl., Kohls, No. 24-cv-3754, 2025 WL 66514, at *4 (D. Minn. Jan. 10, 2025).

[6] *Id.* at *5.

[7] *Id.* at *4.

[8] Order on Joint Discovery. Submissions at 5, Concord Music Group, Inc. v. Anthropic PBC, No. 24-cv-03811-EKL (N.D. Cal. May 23, 2025), Dkt. No. 377.

[9] *Id.* at 6.

[10] Mot. for Leave to Substitute Expert Witness, Expedited Treatment Requested at 3, Khoury et al v. Intermountain Health Care Inc. et al, No. 2:20-cv-00372 (D. Utah Aug. 6, 2025), Dkt. No. 286.

[11] Mot. for Stay at 6, Khoury et al v. Intermountain Health Care Inc. et al, No. 2:20-cv-00372 (D. Utah Aug. 25, 2025), Dkt. No. 292.

[12] Order Dismissing Case at 1, Khoury et al v. Intermountain Health Care Inc. et al, 2:20-cv-00372 (D. Utah Sep. 30, 2025), Dkt. No. 307.

[13] Fed. R. Evid. 702.

[14] *Daubert v. Merrell Dow Pharmaceuticals Inc.*, 509 U.S. 579, 596 (1993).

[15] *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

[16] *Jackson v. Nuvasive, Inc.*, 2025 WL 823229, at *7 (D. Del. Mar. 14, 2025).

[17] Id.

[18] Id. at *8.

[19] Id. at *7.

[20] Id. at *8.

[21] Id.

[22] In re Weber, No. 1845-4/B, slip op. at *11(Sur. Ct., Saratoga County Oct. 10, 2024).

[23] Id. at *12.

[24] Id.