August 23, 2026

AI Claims Under the Microscope: NAD's SafelyYou Decision Shows Why "Proven" Isn't Always Proof

by Bryan J. Wolin

If you're marketing an AI-powered product—especially one that touches health or safety—NAD's recent decision in *SafelyYou, Inc.* (Case No. 7529, closed August 3, 2026) should be required reading.

SafelyYou sells an AI-enabled fall detection and monitoring system for assisted living and memory care communities. Its advertising made a series of bold claims: the system was "proven to reduce falls, risk, and costs while elevating care," delivered "industry-leading capabilities and reliability," and detected falls "with over 99% accuracy." NAD opened the inquiry on its own initiative, as it often does when it discovers health and safety claims that cannot be easily vetted by consumers.

AI Claims Are the New Frontier

This decision underscores what those of us in the ad law world already know: AI-based performance claims are proliferating, and regulators are paying attention. NAD expressly cited another one of its recent decisions for the proposition that in the "emerging AI era," it is "crucial that consumers receive truthful and accurate information concerning this technology." Companies eager to tout “AI-enhanced” offerings need to ensure their claims can actually withstand scrutiny.

Good Data ≠ Supported Claims

Here's what makes this case fascinating: SafelyYou wasn't flying blind. The company submitted an impressive evidentiary package—over 143,000 falls analyzed across nearly 500 communities over five years, peer-reviewed publications, case studies, economic modeling, and satisfaction surveys. And yet NAD still found several claims unsupported or in need of modification.

SafelyYou committed one of the classic advertising blunders: it forgot that even the most robust data cannot save an advertising claim if the data do not fit the claim. The studies it submitted in support of its claim, “Proven to reduce falls, risk, and costs while elevating care,” showed *associations* between system use and improved outcomes, but SafelyYou's own report expressly acknowledged that "no causal attribution" could be drawn. Claiming something is "proven" when your own evidence says otherwise is a recipe for a NAD recommendation you'd rather not receive.

Narrow and Qualified Beats Broad and Bold

The decision also clearly shows how narrow, down-to-earth claims are safer than high-flying ones. SafelyYou's broadest claims—"proven to reduce falls" and "industry-leading"—fell flat. But its narrower, better-qualified claims survived. The claim that its system reduces falls by 40% and ER visits by 80% "in the senior living communities we serve" was found adequately supported precisely *because* SafelyYou limited it to communities actually using the product. The lesson is evergreen but bears repeating: say what your data can support, and nothing else.

NAD Can and Will Do the Work

Some advertisers may assume that in a self-initiated monitoring case—with no adversary marshaling evidence on the other side—NAD won't dig deep into complex, technical substantiation. This decision puts that assumption to rest. NAD methodically evaluated SafelyYou's retrospective analyses, its survey methodology, its economic modeling, and its peer-reviewed literature, distinguishing between claims of correlation and claims of causation with surgical precision. The forum's analytical rigor should not be underestimated, regardless of whether an outside challenger is pushing back.

Health and Safety Claims Demand More

NAD reiterated that claims about health-related and safety-related products require substantiation that "often includes methodologically sound testing with sufficient controls." Translation: if your AI product is being marketed on the basis that it keeps people safer, you'd better have science—real science—behind it.

The Third-Party Evidence Problem in the Age of AI

Finally, a broader point about evidence reliability. Advertisers frequently rely on third-party research to substantiate claims. But a recent pre-print from researchers at Cornell, UC Berkeley, UCLA, and Tsinghua (Zhao, Wang, et al., 2025) documents an alarming trend: LLM-generated hallucinated citations are infiltrating the scientific literature at scale, with an estimated 146,932 fabricated references in 2025 alone across just four

major databases. The contamination isn't limited to a few tainted papers—it's diffusely embedded across many manuscripts.

This paper has not yet undergone peer review, so its conclusions should be taken with a grain of salt. But the Kilpatrick advertising team is monitoring it closely, considering how many advertisers rely on third-party research to substantiate their claims. If and when the paper is formally vetted and published in a reputable journal, we'll do a deeper dive on what it means for advertisers who rely on published literature as substantiation. In the meantime, the message is clear: know your sources, vet your evidence, and don't assume that "published" means "reliable."

For questions about NAD challenges, claim substantiation, or other ad law issues, contact your Kilpatrick Advertising Team.