

News and Perspectives

Classifying Physician Deepfakes: The Many Faces of AI-Powered Medical Deception

Tejas S Athni, JMIR Correspondent

Abstract

Physician deepfakes continue to spread online, with advances in technology leading to more creative—and insidious—methods of deception. In this *News and Perspectives* article, JMIR Correspondent Tejas S Athni reports on how and why different kinds of physician deepfakes can be used to deceive consumers.

Key Takeaways:

- Physician deepfakes are an emerging threat to patient safety and professional trust.
- Physician deepfakes can be categorized into a practical taxonomy with 6 categories based on the primary purpose or means of deception, from existing commercial endorsement, engagement farming, and medical misinformation deepfakes, to fraudulent phishing, potential professional communication, and reputational sabotage.

Tejas S Athni, PhD, is an MD-MBA candidate at Harvard Medical School and Harvard Business School. In this op-ed, he shares a practical taxonomy of the physician deepfakes he has observed or expects to emerge, according to their primary purpose.

Generative AI has reduced the barriers to producing synthetic but realistic impersonations of human behavior and speech. Among the most concerning applications is the rise of physician deepfakes—AI-generated or AI-manipulated content that falsely represents a physician’s identity, likeness, voice, expertise, or professional endorsement—which prey on and [undermine](#) patients’ and society’s trust of the medical profession.

Reflecting the growing concern, the American Medical Association released its first comprehensive policy [frame-work](#) addressing AI-generated physician deepfakes earlier this year. Simultaneously, reports have documented dozens of AI-generated impersonations of physicians across major social media platforms, from Instagram to [TikTok](#). These developments suggest that physician deepfakes are evolving from isolated scams into emerging issues deserving greater attention. How do they deceive and mislead audiences, and to what ends? And what are the potential clinical impacts?

Different Avenues for Deception

Scammers create and deploy physician deepfakes for different reasons and in different forms—with new avenues likely to emerge as the technology continues to evolve. Many of the 6 types of deepfakes described here can both be based on the likeness of a real physician or depict a factitious physician entirely.

Commercial Endorsement Deepfakes

Commercial endorsement deepfakes fabricate a physician’s image, likeness, or voice to make it appear that they endorse a commercial product or company for financial gain. These most often take the form of endorsements for supplements, weight-loss products, or medical devices. There are numerous documented examples involving well-known physicians.

The likenesses of American physicians have been AI-generated to market [vitamins](#). British television physicians have been impersonated in Facebook [advertisements](#) promoting products that claim to cure hypertension and diabetes. AI-generated videos have depicted prominent pediatric academic physicians as recommending [menopause](#) supplements, despite their clinical training and expertise lying outside of menopause care. Deepfake physician images and videos were used to promote [counterfeit](#) GLP-1 weight loss medications through unregulated online pharmacies. Hundreds of [reports](#) have been filed against deepfakes pushing a weight loss supplement called LipoMax.

Across these cases, physician endorsements serve as powerful heuristics that patients frequently use when choosing among health products. Deepfake endorsements falsely leverage years of professional credibility to market interventions that may not have credible clinical evidence.

Engagement Farming Deepfakes

Engagement farming deepfakes are created primarily to generate views, followers, comments, shares, or advertising revenue. They may take the form of Instagram reels, TikTok videos, YouTube personas, or other short-form social media designed to maximize algorithmic engagement. A 2025 CBS News [investigation](#) identified more than 100 videos across TikTok, Instagram, Facebook, and YouTube featur-

ing fictitious or impersonated physicians, with some videos receiving millions of views.

The objective of this variant of deepfake is to create a scalable, recognizable medical [persona](#) that is capable of producing large volumes of attention-generating content across languages and platforms. This exploits the credibility and familiarity associated with physicians, while emphasizing provocative or sensational topics that encourage users to comment and share. The accumulated audience itself generates revenue through views and advertisement clicks, affiliate income, or downstream product sponsorship opportunities. While these deepfakes do not necessarily deceive patients into particular transactions, they manufacture attention at scale by leveraging false medical authority.

Medical Misinformation Deepfakes

Medical misinformation deepfakes are created primarily to present [false](#) medical claims in a legitimate, authoritative way to influence people's health beliefs or behavior. They often combine emotionally persuasive narratives with persecutory or conspiratorial themes; for example, they may suggest that governments or pharmaceutical companies are concealing cures for cancer and chronic conditions, promote [miracle](#) interventions, or encourage viewers to abandon established therapies in favor of [unproven](#) alternatives. AI-generated physicians may also present entirely fabricated medical claims or findings as established fact. More insidiously, this content may fabricate scientific studies in ways that appear authentic by referencing real journals, plausible volume and issue numbers, and legitimate institutional affiliations; this creates the illusion of evidence-based medicine.

Fraudulent Phishing Deepfakes

Fraudulent phishing deepfakes involve using an AI-generated physician to deceive individuals into disclosing protected health information (PHI), financial information (eg, social security number, insurance number), or other sensitive data for downstream exploitation and monetary gain. Rather than merely promoting products, bad actors may increasingly use convincing AI-generated physician impersonations to facilitate scams, including fraudulent requests for payment for nonexistent medical services or medication purchases through [illegitimate](#) online pharmacies. These schemes may be supported by fabricated physician profiles on websites, messaging platforms, or telehealth services.

Professional Communication Deepfakes

Professional communication deepfakes, a theoretical type of future deepfake, could make it appear that a physician is delivering professional, institutional communication—when the individual did not actually create or deliver it—through fabricated continuing medical education (CME) lectures, grand rounds presentations, conference introductions, physician interviews, institutional announcements, or other communications presented as originating from a specific physician. Where authorized synthetic communications involve physicians knowingly permitting reproductions of their likenesses for explicitly disclosed and legitimate professional purposes, this content could create

the impression of an authentic physician-delivered communication, even when the physician was not involved. These deepfakes could be used to mislead audiences, falsely confer medical or institutional authority on a message, influence grant or investment decisions, or solicit research funding or donations.

Reputational Sabotage Deepfakes

Reputational sabotage deepfakes, another theoretical example, could involve manipulations of a real physician's likeness or voice to make them speak or appear in a manner that damages a physician's or institution's professional reputation. They could depict AI physicians making racist, discriminatory, inappropriate, or otherwise objectionable statements or falsely portray them as engaging in professional misconduct, with the objective of damaging a physician's standing among patients, colleagues, employers, or institutions. In competitive settings, such deepfakes could be deployed by competitors or malicious actors to undermine a physician's candidacy for employment or leadership positions or to block professional advancement.

Human Impacts

The harms of physician deepfakes extend beyond misinformation itself. For physicians, unauthorized deepfakes could damage professional reputation, undermine relationships with patients and colleagues, and ultimately threaten livelihood and well-being. For patients, harms could range from financial losses from purchasing ineffective or fraudulent interventions to disclosure and exploitation of PHI. Deepfakes could delay legitimate medical evaluation or encourage patients to pursue dangerous interventions that may cause physical harm.



Preserving the public's trust in physician identity and communication will require sustained professional vigilance.

Looking Forward

Physician deepfakes represent an emerging threat to patient safety and trust in medical expertise. More empirical research should be conducted to get a handle on—and get ahead of—the problem.

Specifically, efforts should be made to validate taxonomies of physician deepfakes, including their primary purpose or mechanism of deception as well as modality (eg, video, voice, static image, avatar), identity target (eg, living physician, deceased physician, fictional persona),

distribution channel (eg, social media, paid advertising, messaging platform, telehealth platform, e-commerce site), and motivation (eg, financial gain, ideological influence, personal malice, benign). Such work can help systematically measure the prevalence and impact of different types of physician deepfakes and ultimately help mitigate the deepfakes posing the greatest risks to patients and physicians.

Clinicians and health systems should develop mechanisms to authenticate physician content, rapidly identify and report fraudulent content, and educate patients about the possibility of synthetic medical impersonation. Preserving the public's trust in physician identity and communication will require sustained professional vigilance.

Keywords: physician deepfakes; AI physicians; generative AI; medical misinformation; health misinformation; AI impersonation; patient safety; medical fraud; physician trust

Please cite as:

Athni TS

Classifying Physician Deepfakes: The Many Faces of AI-Powered Medical Deception

J Med Internet Res 2026;28:e111657

URL: <https://www.jmir.org/2026/1/e111657>

doi: [10.2196/111657](https://doi.org/10.2196/111657)

© JMIR Publications. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 18.Sep.2026