Newcastle upon Tyne
stokel@gmail.com
Cite this as: *BMJ* 2024;386:q1319
http://dx.doi.org/10.1136/bmj.q1319

ARTIFICIAL INTELLIGENCE

Deepfakes and doctors: How people are being fooled by social media scams

ChrisStokel-Walker investigates the increasing prevalence of deepfake videos purporting to be of popular doctors selling scam products

Chris Stokel-Walker *freelance journalist*

Hilary Jones is one of the UK's most recognisable doctors. For decades he's answered the public's questions on television, tackling their worries about illnesses, and making medical jargon easy to understand. It's little wonder, therefore, that people might be interested in the latest wonder drug that Jones endorsed in a video posted on Facebook earlier this year.

The drug was purported to be a cure for high blood pressure, and Jones was supposedly talking about it on the *Lorraine* programme, on which he often appears.[1]

The video was, however, not of Jones. It was a deepfake created by artificial intelligence (AI) technology. And it's far from the only one.

"Some of the products that are currently being promoted using my name include those that claim to fix blood pressure and diabetes, along with hemp gummies with names like Via Hemp Gummies, Bouncy Nutrition, and Eco Health," says Jones.

He isn't alone in seeing his name co-opted. Michael Moseley and Rangan Chatterjee are just two of the other doctors whose public profiles have been used to promote health scams.

Seeking out scams

John Cormack, a retired doctor based in South Woodham Ferrers in Essex, worked with *The BMJ* to investigate the scale of deepfake doctors across social media.

"Previously I steered away from things like Facebook but during the pandemic it became obvious that a lot of misinformation was being shared online," he says. Cormack acted directly by posting about misinformation that his patients might encounter. But sharing posts on local Facebook pages and writing columns for his local news magazine could only reach patients in his own area.

After the acute phase of the pandemic passed, he saw a trend towards deepfake videos. "The latest trend is deepfake, where doctors who are perfectly innocent get embroiled in ideas such as big pharma and the health services withholding the cure for diabetes—and if you take these expensive pills for a short period of time it can cure you,'" he says.

Deepfakes use AI to map a digital likeness of a real human onto a video of a body that isn't theirs. The technology—at least in its current state—is relatively new, and so reliable evidence on how convincing it is can be hard to come by. But one recent study of deepfakes, specifically talking about scientific subjects, suggests that 25-50% of people cannot distinguish them from authentic videos.[2]

"The bottom line is, it's much cheaper to spend your cash on making videos than it is on doing research and coming up with new products and getting them to market in the conventional way," says Cormack. "They seem to have found a way of printing money."

He believes the General Medical Council would see it as outside their jurisdiction. "The GMC would only do something if the videos were genuine" and doctors were promoting bad products, Cormack says. "It sees its role as more of a regulatory body."

Jones also doesn't believe it's the GMC's job to advocate for doctors in battles against social media companies. "While it would be aware of the scams, I'm not sure what means it has to do anything about them," he says.

A GMC spokesperson told *The BMJ*, "The GMC's regulatory powers, as set out in the Medical Act, only extend to individual doctors who are on the UK medical register. Computer generated videos by people not on our register would sit beyond our remit." The spokesperson said the organisation sympathised with doctors.

Because of that, the only way Cormack sees of tackling the problem is to make the platforms that host the content—Facebook, Instagram, X (formerly Twitter), YouTube, and TikTok, to name a few—accountable for the videos they help disseminate.

"It's down to the likes of Meta, the company that owns Facebook and Instagram, to stop this happening," says Jones. "But they've got no interest in doing so while they're making money."

A Meta spokesperson told *The BMJ*, "We will be investigating the examples highlighted by *The BMJ*. We don't permit content that intentionally deceives or seeks to defraud others, and we're constantly working to improve detection and enforcement. We encourage anyone who sees content that might violate our policies to report it so we can investigate and act."

The tech behind the con

The slew of questionable content on social media that uses the likenesses of popular doctors and celebrities is an inevitable consequence of the AI revolution we're currently living through, says Henry Ajder, an expert on deepfake technology. "The rapid democratisation of AI tools for voice cloning and avatar generation has transformed the fraud and impersonation landscape," he says.

The ubiquity of these tools means that they've been taken out of the hands of specialists and put into the hands of the general public—where they can create a critical mass of content that bombards users. "There's been a big increase in this kind of activity," says Jones, who employs a social media specialist to trawl the web for deepfake videos that misrepresent his views and tries to have them removed. "Even if they're taken down, they just pop up the next day under a different name," he says.

"Some of these tools require identity checks or biometric authorisation, but many do not have robust safety measures to stop bad actors puppeteering someone's likeness," Ajder says. "Over the past year we've seen huge growth in this form of deepfake fraud, particularly on YouTube and X. Many are selling fraudulent cryptocurrencies, investment schemes, or medical products, with varying degrees of sophistication."

Deepfakes work by preying on people's emotions. They use the image, voice, or likeness of a trusted person in a community. The same thinking was the reason one AI company, OpenAI (whose products are not believed to have been used in any of the videos seen by *The BMJ*), sought to convince Hollywood actor Scarlett Johansson to voice its ChatGPT tool, GPT-4o, released in May 2024. Johansson refused but OpenAI used a similar sounding voice—at which point Johansson complained. OpenAI quickly took the ability to use the voice in question offline.

An emotional connection

When it comes to medical products, that emotional connection with the person telling you about the new wonder drug or magnificent medical device matters even more. Someone you don't know trying to sell you the virtues of a particular treatment may raise suspicions. But if they're someone you "know" from social media, television, or radio, you're more likely to believe what they're saying.

Indeed, it's the reason that doctors like Jones take to the airwaves—he has previously said that he's able to influence tens of thousands of people with a single TV appearance in the same time it would take him to speak to one patient in a clinic.[3]

Spotting deepfakes can be tricky too, says Ajder, as the technology has improved. As recently as a year ago, there were telltale signs that a video was AI generated. Sometimes you could see multiple fingers or blurry sleeves; at one point, a near certain giveaway was that AI tools couldn't correctly render human earlobes. In other instances, a disconnect between what the person was saying, and how their mouth moved, could be discerned.

"Some combine synthetic voice audio with an existing video where lip movements don't match perfectly, whereas others combined voice cloning with synthetic lip synchronisation techniques," says Ajder. "It's difficult to quantify how effective this new form of deepfake fraud is, but the volume of videos now circulating would suggest bad actors are having some success."

For those whose likenesses are being used, there's seemingly little they can do about it (see box). Jones says, "The onus falls on people using social media not to buy anything from it, because it's so unreliable that you simply don't know what you're buying."

What to do if you find a deepfake

- Firstly, look carefully at the video or image, or listen to the audio, to make sure your suspicions are well founded. In a post-AI era we've all become more sceptical and there can be a risk of becoming "the boy who cried wolf"
- Try to contact the person endorsing the product if you aren't the subject of the video, and see if the video, image, or audio is legitimate. Do so using a method away from where you encounter the video—it's possible the social media account with the questionable content is also controlled by a scammer
- Leave a comment on the content questioning its veracity. You may not be the only person with concerns and letting people know that you think all is not right may help dissuade others from parting with their cash
- Use the platform's built-in reporting tools. You will often find these alongside the post, usually hidden behind three dots which are frequently in the top right corner of the post. You may be asked to share more information about your concerns and the report will be sent through an automated system. If platforms also believe it is fake, they will remove it
- Also report the person or account that shared the post. Simply removing a single post may not be enough and sometimes social media systems aren't intelligent enough to recognise repeat offenders
- Even if you're not the subject of a deepfake, you should be aware that patients may come to you having seen deepfake videos. They may want certain treatments deepfakes have recommended. You should try to make them aware that deepfakes exist, and that they ought to follow standard medical advice rather than what they have seen on social media

Competing interests: None declared.

Commissioned; externally peer reviewed.

1 Good Morning Blog. Hilary Jones: the main enemy of high blood pressure has been found. Facebook. www.facebook.com/reel/787551762897006

2 Doss C, Mondschein J, Shu D, etal. Deepfakes and scientific knowledge dissemination. *Sci Rep* 2023;13:. doi: 10.1038/s41598-023-39944-3 pmid: 37596384

3 Scott J. TV doctors: taking medicine to the small screen. BBC News. 28 August 2018. www.bbc.com/news/health-45295982