

AI in the NHS and reputational risk

By [Jack Myers](#), Associate Director at reputation, risk and change advisory firm, [Alder](#)



This month, complaints about a recently installed AI receptionist, “Emma”, in Rotherham were widely reported. One patient told the media: “One of the issues is this system can’t always understand what people’s inquiry is about due to their broad Yorkshire accent.” While the supplier said patients could request a human at any point, some described giving up on Emma entirely in exasperation.

It may prove to be a relatively minor hiccup that is smoothed out, but it points to a growing area of reputational risk for NHS organisations: What happens when the public

care, compared with 54% of the public.

The scale of adoption is also changing rapidly.

NHS England announced in June that more than 500,000 clinicians and support staff would receive access to Microsoft 365 Copilot, following a trial involving more than 30,000 staff across 90 NHS organisations.

For a trust, the scenarios are easy to imagine: a fake video of a consultant making a controversial clinical claim...

does not understand or accept how AI is being used?

AI is already being adopted across the NHS at pace, and there is a visible confidence gap. [Health Foundation polling published this year](#) found that 80% of NHS staff supported AI being used in patient

That trial reported average time savings of 43 minutes per staff member each day, and its proposed uses range from drafting patient letters and supporting discharge processes to rota and bed management, HR, procurement and preparation of board papers.

there is an important distinction between different types of technology. [New NHS England and MHRA guidance](#) makes clear that a product which simply transcribes or summarises a consultation may not be regulated as a medical device.

A system that supports diagnosis, treatment or automated clinical actions, however, may fall within medical-device regulation. That makes understanding and

normally required before an ambient scribe is used.

But, because the technology is new, NHS guidance says patients should currently be told at the beginning of the consultation that it is operating and given the opportunity to object or dissent.

If they do not want it used, it cannot be used.

Before deployment, then, trusts need to be able to explain in plain English what the technology does;

Crisis scenario training should be updated to add this new layer of AI complexity.

communicating exactly what a particular system does essential.

It also creates a risk of function creep.

A product introduced as an administrative tool may acquire additional capabilities as it is updated, and NHS England therefore expects organisations to monitor changes in functionality and consider whether those changes alter the product's regulatory status or risk profile.

For individual care, explicit consent is not

- whether the consultation is recorded;
- where any audio or transcript is stored;
- how long it is retained; when it is deleted;
- whether a commercial supplier holds a copy or can use the information for any other purpose; and
- who remains accountable for the final clinical record.

A system can be technically compliant and clinically useful and still suffer a reputational failure if patients first discover what it does when they encounter it in



the consulting room or through a critical newspaper story.

This also means communications teams should be brought into AI deployment earlier.

NHS guidance envisages CIOs, clinical information leads, SIROs, Caldicott Guardians, DPOs and boards being involved in decisions around ambient technologies.

Communications should sit alongside them.

AI also creates reputational risk from outside the organisation. Deepfake audio and video can convincingly imitate trusted individuals. The BMJ has documented fake videos of well-known UK doctors apparently endorsing treatments and products.

In January 2026, the Government acknowledged the threat of AI-generated deepfakes impersonating medical professionals to spread health misinformation.

For a trust, the scenarios are easy to imagine: a fake video of a consultant making a controversial clinical claim; a cloned executive voice announcing a service closure or data

breach; a manipulated patient video alleging mistreatment; or an apparent member of staff seeking confidential information.

Crisis scenario training should be updated to add this new layer of AI complexity.

In terms of preparation, though, the first task is establishing the facts of the organisation's own usage and systems:

- knowing where AI is being used across the trust;
- understanding what each system is actually capable of doing;
- identifying who is accountable for checking outputs...

... and ensuring communications, legal, clinical, cyber and information-governance teams know how they would respond to an AI-related incident.
