BRIEFING · SEPTEMBER 2026

Artificial Intelligence in *UK HealthTech*

What the evidence actually says, what the regulator expects, and what can safely be done now.

Neil Catton

Independent Technology Strategist · Technology, Ethics & Human Purpose

NEILCATTON.COM

Prepared as a market briefing. Every figure is traced to a regulator, a trade body, an ombudsman, published legislation, or an operator speaking on its own record.

ONE

A note on the numbers

Almost every published statistic about AI in health is produced by a company selling AI into health. None of them appear here.

This briefing is built from a narrow set of sources, deliberately. The MHRA, in its guidance, its regulatory sandbox reports and its inspectorate. NHS England, in its guidance, its priority notifications and its own trial data. An NHS foundation trust evaluating the technology on its own staff. The Nuffield Trust with the Royal College of General Practitioners. The Office for Life Sciences. A university research team that published its method and its funder. Where a figure could not be traced to one of those, it is not here.

That discipline costs something. There is no market-size figure for AI in UK health technology in this document, because every one located traced back to a commercial research house behind a paywall with no published method. There is no supplier accuracy claim, no time-saving claim from a vendor, and no adoption statistic from a trade survey whose edition could not be established. Naming what was left out is more useful than any of it would have been.

What was excluded, and why

- → **Every supplier accuracy and time-saving claim.** The regulator and two NHS trusts provide better figures from parties with nothing to sell.
- → **All market sizing for UK health AI.** No published methodology could be found behind any figure in circulation.
- → **A widely circulated trade body adoption survey.** Undated in its published form; the edition could not be established, and an undated recurring survey is a date trap.
- → **A regulator's commentary on AI adoption** that reads as current, ranks highly in search, and is dated June 2022 — its stated plans were for 2022 and 2023.
- → **A trade-press figure of 20,000 clinicians** in a London rollout, superseded by NHS England's own published numbers.
- → **Outcome figures for the NHS AI Diagnostic Fund.** Widely quoted; no published evaluation located.
- → **Any claim that clinicians are failing to report incidents.** As section six sets out, the reporting route does not exist for most of these products. The finding is structural, and stating it as a behavioural failure would have been both wrong and unfair.

Read but not used

The CQC's Artificial intelligence in health and social care: CQC's role, expectations and plans (21 May 2026) could not be retrieved directly, and no position is attributed to the CQC here on the strength of secondary accounts. A CQC commentary on AI adoption ranking highly in search was found to date from June 2022 and is excluded.

One figure is used with its limits stated

The MHRA's finding that out-of-scope behaviour was detected in approximately 39% of real-world notes carries no stated sample size in the source document. It is used here because it is the regulator's own measurement of the central risk, and its missing denominator is reported alongside it rather than concealed. Where two credible surveys disagree — and two surveys of GP AI use do — both are given, the better-sampled one leads, and the disagreement is named.

Sources for every section are given at the foot of that section and listed in full in section thirteen.

TWO

The sector this lands on

UK medical technology is the larger half of the life sciences sector by company count and by employment, and it is now the object of the largest technology commitment the NHS has made.

4,360
medical technology companies — 60% of the UK life sciences sector

196,000
people employed in UK medical technology — 55% of the sector

£48.0bn
medical technology turnover, against £98.9bn for biopharmaceuticals

Those are the Office for Life Sciences figures for 2023/24, published in October 2025 and the most recent available. The publication is candid about their limits, and that candour belongs in any document that uses them: they are designated "official statistics in development", they are not comparable with 2021/2022 or earlier because of a methodology change, and the department states that "limited information is known about its coverage and what the true population of the life sciences sector is in the UK". They are the best figures there are. They are not precise, and this briefing does not treat them as though they were.

What is being spent

On 4 July 2026 NHS England announced £10bn over three years for technology, digital and data systems, with stated benefits of £41bn over the decade. The announcement named an AI triage tool in the NHS App reaching more than 200,000 patients within twelve months and every NHS App user by April 2028; ambient voice technology expanding across four south-west London trusts and to more than 3,000 clinicians at Alder Hey and Manchester University NHS Foundation Trust; and Microsoft Copilot access for more than 500,000 NHS staff.

> THE WARRANT FOR THE LARGEST ELEMENT
>
> Sir Jim Mackey, Chief Executive of NHS England, described clinicians gaining "up to a quarter more of their time with patients". That figure comes from a specific study, of a specific product, across nine London sites. Section three examines what that study measured, and what it did not.

Office for Life Sciences, *Bioscience and health technology sector statistics 2023 to 2024*, published 2 October 2025. NHS England, "NHS accelerates artificial intelligence rollout to cut waiting times and improve care for millions", 4 July 2026.

THREE

The pressure point

Two arms of the state evaluated the same product and produced two different pictures. The rollout followed one of them.

In February 2026 the Data Research, Innovation and Virtual Environments unit at Great Ormond Street Hospital published the final report of an NHS England-sponsored evaluation of ambient voice technology. It is a serious piece of work: a non-randomised multi-centre within-subject pre-post trial across nine sites, 192 clinicians and 17,452 patient encounters, using TORTUS from Tortus AI. It is published in full, its limitations are declared, and its findings are strong.

WHAT IT MEASURED

Direct care time rose from a median of 70.0% to 86.5% — a 23.5% relative increase, across 104 clinicians and 2,095 observations. Consultation length fell 8.15%. Documentation time fell 51.7% at individual level. Emergency department capacity rose 13.4% per shift. Cognitive load fell significantly on a validated instrument.

WHAT IT DID NOT MEASURE

Transcription errors. Hallucinations. Omissions. Clinical inaccuracy. None of these is measured or systematically discussed anywhere in the report. It also states that a documentation quality instrument used in an earlier phase was dropped, and that the team "decided not to score the documentation quality due to lack of objectivity".

The report gives its reason plainly, and it is a reasonable one: "In light of potential consequences arising from hallucinations and omissions and to mitigate against their impact, the technology we tested is designed to work with human (clinical) oversight and require approval of the output as part of a clinician's professional responsibility." The safety case is human review. The study therefore measured what human review delivers — time, capacity, experience — and not what it is there to catch.

The same product, in the regulator's sandbox

In June 2026 the MHRA published the programme report for phase two of AI Airlock, its regulatory sandbox. Seven products were selected from 51 applicants and tested between April 2025 and March 2026. TORTUS was one of them.

THE MHRA'S FINDING

"Without active guardrails, out-of-scope performance was detected in approximately 39% of real-world notes, falling to 20% with guardrails active."

Out-of-scope, in this context, means the system moving from documentation into clinical decision support — doing something its intended purpose does not cover. The report gives no sample size for that finding, and this briefing does not supply one. It also records, in the same passage, a real-world precision score of 0.989 alongside the observation that "users may reduce scrutiny of outputs over time in response to consistent and reliable product performance". A product that is nearly always right is a product that stops being checked.

GOSH DRIVE, *The use of Ambient Voice Technology with Generative Artificial Intelligence in Multiple Clinical Settings Across the NHS*, version 1.3, updated 11 February 2026. MHRA, *AI Airlock Sandbox Phase 2 Programme Report*, June 2026.

FOUR

What the sector is facing

Where error figures do exist, they come from a single trust that chose to publish them, and from a national survey of the clinicians using the technology.

One trust, four products, 9,239 encounters

TheHill at Oxford University Hospitals NHS Foundation Trust evaluated four ambient voice products — Accurx Scribe, Heidi, Lyrebird and TORTUS — across 136 users and 9,239 encounters, and published in February 2026. It measured what the larger study did not.

44.4% of participants encountered situations in which the technology invented information

6.2% of logged uses recorded a hallucination — 24 of 398. The two figures measure different things: one asks whether a clinician has ever seen it, the other whether it happened on that occasion

37.3% of outputs required editing — 126 for clinical accuracy, 43 for spelling and grammar, 24 for hallucinations

81.5% agreed the record produced was accurate — the benefit and the risk sit in the same dataset

Performance degraded where the consultation was hardest. In Community Paediatrics, 60% of users required extensive editing and 40% reported frequent confabulation or omission. Critical Care showed the same pattern. Multi-voice ward rounds and qualitative documentation performed poorly. The trust's own limitations section is worth quoting because it is more honest than most: the data is "less robust for drawing definitive, quantitative conclusions about AVT's impact on objective outcomes such as time savings, error rates, patient safety, or clinical effectiveness".

What clinicians are actually doing

The Nuffield Trust, with the Royal College of General Practitioners, surveyed 2,108 practising GPs between 28 July and 20 August 2025, achieving a 28% response rate from a membership frame of around 50,270, alongside five focus groups. It found 28% of GPs using AI tools in clinical practice — 13% with practice-provided tools, 11% with tools they obtained themselves, and 4% with both. Among those using AI, 57% use it for clinical documentation and ambient voice, and 28% for clinical decision support. Only 5% had received any formal training.

A separate survey published in *npj Digital Medicine* in May 2026, of 598 GPs recruited through a commercial panel, found 40% currently using AI scribes across 34 different products, applied to a mean 60% of consultations. The two surveys disagree. They used different frames, different questions and different recruitment, and the Nuffield Trust study has the stronger sample. This briefing leads on 28% and reports the disagreement rather than choosing the larger number.

Where the gap is not documentation

Every figure above concerns one category of AI: the kind that writes down what a clinician says. It is not the only category in use, and it is not the fastest-growing. A national imaging programme is expanding faster than the evidence behind it, an autonomous diagnostic tool is already discharging patients without a doctor seeing the image, and a regulatory sandbox is testing four more tools this briefing had not previously used.

THE NATIONAL ROLLOUT

Imaging AI, funded to reach every trust by 2029

GOV.UK announced a further £30 million for AI across the NHS on 10 June 2026 — £20 million to roll out AI X-ray tools to every NHS trust in England by 2029, plus £8.1 million to pilot AI on CT, ECG and X-ray data across thirteen sites, for heart failure, stroke, lung cancer, lung infection and tic disorders. Roughly half of England's NHS trusts already have some form of this technology. Over 4 million patients have received a faster lung cancer diagnosis or all-clear from it. The NHS performs over 7 million chest X-rays a year, and average scan analysis time is now around 4 days against 8 days previously for the most complex cases.

THE ONLY INDEPENDENT EVALUATION SO FAR

Modest gains, no proven patient benefit yet

An NIHR-funded evaluation by the Rapid Service Evaluation Team, published 16 March 2026, covered 66 NHS hospital organisations using AI for chest diagnostics under the Fund, ten examined in depth. AI correctly prioritised likely-cancerous scans, but turnaround-time improvements were "modest" and varied by site, follow-up referral rates moved in opposite directions at different sites, and — the finding that matters most — overall impact on patient outcomes was assessed as negligible so far, despite the tools being judged cost-effective. Many patients did not know AI had been used in their care, despite wanting to be told.

THE CLOSEST THING TO AUTONOMOUS, LIVE TODAY

A cancer pathway that can discharge without a doctor

Skin Analytics' DERM holds the UK's highest device risk classification, Class III, and NICE gave it a conditional recommendation. It runs a fixed algorithm against dermoscopic images and can autonomously discharge patients it assesses as benign from the urgent suspected-skin-cancer pathway — over 650,000 referrals a year — without a clinician reviewing the image. NHS England's own implementation guidance requires informed consent, a clinician-reviewed run-in period, and a mandatory second clinical read for every patient with black or brown skin for the whole three-year evidence-gathering period, because the evidence base for accuracy on darker skin tones was judged insufficient to rely on the algorithm alone.

MHRA's AI Airlock sandbox is testing seven tools through to March 2026, four of them pattern-prediction from data with no documentation angle at all: Panakeia's PANProfiler reads pathology slides with AI to call a genetic biomarker status that would otherwise need a separate test; Octopath is a general AI pathology platform; Eye2Gene infers a likely genetic diagnosis from a retinal image; and DeepX Health's DeepX AI interprets blood test results. None of the four appears anywhere else in this briefing.

Two further developments sit underneath all of this as infrastructure rather than as deployed AI. The NHS Federated Data Platform — a £330 million, seven-year, Palantir-led contract covering up to 240 NHS organisations — pools operational and clinical data so integrated care boards can run population health management tools; NHS England's own account describes it as "a foundation for future innovation, including AI tools" rather than a predictive system in its own right today. And a national demand-forecasting tool, in use by 50 NHS organisations since December 2025, already predicts daily accident and emergency demand from Met Office weather data and admissions history — a genuinely operational example of prediction from health-system data, distinct from clinical diagnosis. Separately, Our Future Health — 2.4 million UK adults consented toward a target of 5 million, combining genetic, NHS record and lifestyle data — is the population resource this generation of predictive and preventative research increasingly depends on.

TheHill at Oxford University Hospitals NHS Foundation Trust, *Ambient Voice Technology in Acute Services — An Evaluation*, February 2026. Nuffield Trust and RCGP, *How are GPs using AI? Insights from the front line*, December 2025. Derksen et al., *npj Digital Medicine*, 16 May 2026, funded by Barts Charity and Cancer Research UK. GOV.UK, *AI to speed up cancer diagnosis for millions of NHS patients*, 10 June 2026. NIHR RSET, rapid evaluation of AI in chest diagnostics, 16 March 2026. NHS England, AI-based skin lesion analysis technology (DERM) implementation guidance. MHRA, AI Airlock Phase 2 tool list, 16 October 2025. NHS England, Federated Data Platform FAQs. DSIT/GOV.UK, AI demand-forecasting tool for A&E, 28 December 2025. Our Future Health, cohort profile, International Journal of Epidemiology, 15 October 2025.

FIVE

The challenges of using AI here

Five risks, each named by a regulator, the sector's own trade body or the statute book rather than by this author.

ONE — OVERSIGHT DECAYS

The better it gets, the less it is checked

The MHRA recorded a precision score of 0.989 and, in the same paragraph, that users "may reduce scrutiny of outputs over time in response to consistent and reliable product performance". Human review is the entire safety case for every product outside device regulation. It is also the part that erodes silently, is never measured, and cannot be reconstructed after the fact.

TWO — DRIFT WITHOUT CHANGE

The product changes when nobody changed it

"Device behaviour can shift significantly without explicit design changes." A validation performed at procurement describes a system that may no longer exist. Nothing in current assurance requires anyone to look again, and the AI Airlock report notes that the product under test was continuously improved during the trial.

THREE — UNEVEN PERFORMANCE

Acceptable overall, worse for some patients

AI devices "may perform acceptably against overall performance metrics while underperforming for specific patient subgroups", and those disparities "may be invisible at pre-market evidence but become apparent after deployment". One AI Airlock participant placed underrepresented skin tones at the centre of its investigation. In ambient voice, accent, first language and speech difference are the equivalent, and the Oxford evaluation names transcription difficulty with accents and speech disorders among its risks.

FOUR — COMPETITIVE PRESSURE ON THE CAREFUL

The conservative supplier loses the tender

The trade body's own words, submitted to the MHRA in February 2026: companies taking a conservative regulatory approach "can then find themselves competing in tenders against products making similar claims without the same level of device oversight. That is corrosive to trust and investment." It also observes that "in class I/self-declared territories, enforcement has been especially low-visibility".

FIVE — READ THE WRONG WAY ROUND

Section 80 did not liberalise health data

From 5 February 2026, UK GDPR Articles 22A–22D open the full range of lawful bases for most significant automated decisions, subject to four safeguards. Section 80 is widely described as having made automated decision-making permissible. In this sector that is misleading: the prohibition was narrowed, not removed, and what it still catches is the decision based entirely or partly on **special category data, which includes health data**. Those decisions still require an Article 9 basis such as explicit consent.

The weight therefore rests, as before, on whether the human review was real — and nothing obliges anyone to measure whether that review is still happening at month six.

MHRA, *AI Airlock Phase 2 Programme Report*, June 2026. ABHI response to the MHRA call for evidence, February 2026. TheHill at Oxford University Hospitals, February 2026. DUAA 2025, s.80.

SIX

The threat side

In this sector the threat is not something done to the organisation. It is the entirely legitimate use of unregulated tools by patients who cannot get seen — and, separately, the fact that nothing that goes wrong can currently be counted.

The record nobody can keep

This is the most consequential finding in the briefing and it needs stating carefully, because it is not a criticism of clinicians or of the organisations deploying the technology.

When the MHRA confirmed on 29 July 2026 that products solely intended to transcribe, summarise, draft correspondence or suggest clinical codes are not medical devices, it resolved a genuine ambiguity, in a direction the trade body had asked for. It also placed the most widely deployed clinical AI in Britain outside the Yellow Card scheme — the national route by which a fault in a medical device is reported and counted.

THE REPORTING GAP, STATED BY A PATIENT SAFETY CHARITY

"Currently there is no reliable mechanism to detect, report or count AVT-related patient safety incidents."

Non-medical-device products are not reportable through Yellow Card, and NHS England's Learn from Patient Safety Events service "has no established route for recording that an incident involved AI-generated documentation".

When the MHRA was asked in August 2025, a Yellow Card search had identified no adverse incident reports relating to AI scribes at all. That figure is often read as reassurance, or as under-reporting by clinicians. It is neither. The National Commission's own findings list "limited awareness that Yellow Card scheme applies to AI-enabled medical devices" among the gaps — and after July 2026 most of these products are not medical devices in any case.

The consequence for a board is precise. An organisation deploying this technology today cannot demonstrate its safety record, because no mechanism exists through which one could be built. Neither can it be shown to be unsafe. A board asked how many times the system has been wrong in its own service has no answer available to it, and will not have one until it builds the count itself.

The clinical AI that is not yours

In May 2026 King's College London's Policy Institute, with King's Health Partners and Responsible AI UK, published survey work conducted by Focaldata among 2,093 UK adults, fielded 24–30 March 2026 and weighted by age, gender, region and 2024 general election vote.

15%

used an AI chatbot for health advice instead of contacting a GP or other NHS service

21%

of those who used AI for health advice decided against seeking professional care because of it

6%

would hold the developer responsible if AI missed a problem. 34% would hold the clinician

The report also finds that "the public significantly overestimate how widespread either use of AI currently is among GPs, but particularly in clinical decision-making, where the public's average guess is five times greater than the reality" — 39% against an actual 8%, the latter from the 2025 Nuffield Trust and RCGP survey.

None of this is misuse and this briefing does not treat it as such. It is what people do when the regulated route is slow. The operational consequence is specific: a patient may arrive already triaged by a system with no oversight, no audit trail and no liability — and if something is then missed, a third of the public would hold the clinician in front of them responsible.

MHRA guidance, 29 July 2026. Patient Safety Learning, *Ambient voice technology: exploring the patient safety risks*, 25 August 2026. MHRA statement to *Pulse*, 6 August 2025. Clark, Duffy, Kim, MacAskill, May and McCann, *The use of AI in UK healthcare: public perceptions and healthcare priorities*, King's College London, May 2026.

SEVEN

The UK legal position

Two public bodies gave opposite answers to the same question thirteen months apart, and neither document says what happens to the products deployed in between.

9 JUNE 2025 — NHS ENGLAND

Summarisation is a medical device

A priority notification from Alec Price-Forbes, National Chief Clinical Information Officer, stated that "all AVT solutions that undertake summarisation require, at least, MHRA Class 1 medical device status", and directed organisations to "pause, reject or stop engagement with any AVT supplier that is not able to meet the published assurance standards".

29 JULY 2026 — MHRA

Summarisation is not a medical device

Products solely intended to transcribe and summarise clinical conversations, draft correspondence or suggest clinical codes for clinician review "are not to be regulated as medical devices under the current regulatory framework". NHS England published an aligned position the same day. No transition period is stated.

In January 2026, between those two dates, NHS England published an ambient voice supplier registry listing nineteen suppliers assessed against criteria that required Class 1 medical device accreditation and a current Digital Technology Assessment Criteria assessment. Those suppliers qualified against a requirement the MHRA has since said does not apply. NHS England's July 2026 position addresses prospective procurement and says nothing about products already deployed under the earlier rule.

The sentence that has not been withdrawn

> *"Liability for the use of non-compliant AVT solutions will be held by the local NHS trust, primary care practice or individual clinicians."*

Alec Price-Forbes, National Chief Clinical Information Officer, NHS England, 9 June 2025. To *HSJ* ten days later he added that "proceeding with non-compliant solutions risks clinical safety, data protection breaches, financial exposure, and fragmentation of broader NHS digital strategy".

And a second body of law that applies to all of it

Everything above concerns whether a product is a medical device. Every one of these systems also processes personal data, most of it health data, so a second regime applies regardless — including to the products the MHRA has just placed outside device regulation. It moved twice in 2026, in opposite directions. Section five sets out what that means for automated decisions; section ten sets out the code of practice now being written under it.

NHS England priority notification, 9 June 2025. Ambient voice registry, 19 January 2026. MHRA guidance and NHS England, *Medical device regulation for ambient voice technology products*, both 29 July 2026. Data (Use and Access) Act 2025, s.80, in force 5 February 2026. SI 2026/425, in force 12 May 2026. ICO, *AI and biometrics strategy update*, March 2026.

EIGHT

Beyond the UK

The framework that would answer most of the questions in this briefing was due in the summer and has not arrived.

The National Commission

The independent National Commission into the Regulation of AI in Healthcare was established to develop regulatory recommendations for AI use in healthcare. It is chaired by Professor Alastair Denniston, with Professor Henrietta Hughes, Patient Safety Commissioner for England, as deputy chair, and works through four groups led by the MHRA's Chief Executive, Professor Cathie Sudlow, Professor Hughes and Professor Neil Lawrence.

Its call for evidence closed on 2 February 2026 and drew 761 responses — 511 individuals and 238 organisations. The findings were published on 11 June 2026. Its recommendations were due in summer 2026. As at the date of this briefing they have not been published; the gov.uk page was last updated on 19 June 2026 and lists eight sets of meeting minutes, the most recent from 13 May 2026.

81%	of patients and the public said safety and performance standards are insufficient — 46% strongly
48%	of industry respondents said data governance is already sufficient, against 82% of the public who said it is not
77%	of patients and the public said the framework is somewhat or too loose. 47% of industry called it somewhat restrictive
77–88%	across every group said the liability framework is insufficient or has gaps. Only 14% of industry called it sufficient

One response, from a public body, states the problem this briefing is about better than any summary: "Responsibility for safe AI use should be shared, explicit, and proportionate, rather than implicitly pushed onto frontline clinicians." Set against the June 2025 liability sentence, that is the whole argument in two quotations.

Two comparators

Ireland. HIQA has published national guidance for the responsible and safe use of artificial intelligence in health and social care services, developed through public consultation and a co-production working group. Whatever its merits, a comparable UK document does not exist.

The European Union. AI in medical devices falls under the EU AI Act's high-risk regime rather than its prohibitions, with the substantive obligations deferred. Organisations supplying into the EU should verify the current date directly rather than relying on any secondary account, including this one — it has already moved once.

The UK is not as empty as the comparison suggests: the gap is health-specific guidance, not AI regulation as such. Section seven sets out the statutory AI code now being written by the Information Commissioner.

For a UK provider, the absence of a national framework is not freedom. It is the obligation to construct a defensible position from first principles and hold it, with no published standard to point at when asked. For a UK supplier selling abroad it is harder still: the buyer's regulator will have a framework, and the product will be assessed against it whether or not it was built for one.

National Commission into the Regulation of AI in Healthcare, gov.uk, page last updated 19 June 2026. MHRA, *call for evidence summary of findings*, 11 June 2026. HIQA, *National Guidance for the Responsible and Safe Use of Artificial Intelligence in Health and Social Care Services*. EU AI Act, Annex III.

NINE

Today

Seven things that can be done now. None requires the National Commission to report first, and all of them are stronger for being done before it does.

01	**Build the count nobody else is keeping**	Instrument your own deployment. Log every correction a clinician makes to an AI-generated note, classify it as clinical, factual or presentational, and report the rate monthly. Given the gap in section six, it is the only safety record that will exist for your service.
02	**Test the review, not the product**	The safety case rests entirely on clinicians reading the output. Measure whether they still do at month six as closely as at week one. Automation bias is measurable — sample notes, re-read them independently, and compare. It is currently measured nowhere.
03	**Re-check device status against use, not procurement**	Around a tenth of scribe users apply them to clinical decision support. That use makes a product a medical device regardless of how it was bought or what the supplier's intended purpose says. Audit what staff actually do with it, and act on the answer.
04	**Fix the incident route locally**	If the national system has no field for AI involvement, add one to your own reporting. Otherwise the incidents that do occur will be recorded as documentation errors with no attributable cause, and the pattern will be invisible to you as well as to the regulator.
05	**Test performance across your own population**	Subgroup disparities are invisible pre-market and appear after deployment. Accent, first language and speech difference are the ones to sample deliberately, and the Oxford evaluation shows the effect is largest in the most complex consultations.
06	**Train the five per cent up**	Formal training reaches almost nobody. Automation bias, hallucination patterns and the limits of a product's intended purpose are teachable in an hour, and are currently taught to almost no one who is accountable for the output.
07	**Write down who is accountable, before you are asked**	NHS England put liability on the trust, the practice and the individual clinician in writing in June 2025 and has not withdrawn it. Decide internally who carries what — the clinician, the clinical safety officer, the board — and record the decision now rather than after an incident.

None of the seven requires supplier engagement, new procurement, or the outcome of any pending regulatory process.

Derived from the evidence in sections three to seven.

TEN

Tomorrow

Three levels, defined by where accountability sits rather than by how capable the system is. That distinction is what makes the terms useful to a board.

LEVEL ONE

Assistive — the person decides

Prepares, summarises, surfaces. Accountability is unchanged. Ambient scribing sits here, and the entire regulatory settlement of July 2026 depends on it staying here. Almost all defensible value in this sector is at this level today — and the MHRA's 39% finding describes products drifting out of it when nothing holds them in.

LEVEL TWO

Augmentive — the person confirms or overrides

Recommends within bounds, with reasoning and confidence, and the override is recorded. The review must be demonstrably meaningful, which since 5 February 2026 is a legal requirement rather than a design preference. Most triage and decision-support products belong at this level and are frequently procured as though they belonged at the first.

LEVEL THREE

Adaptive — the envelope is what is overseen

Decides within a defined envelope. Oversight moves to the envelope; accountability moves to whoever sets and monitors it. This is a reasoned projection and is offered as one. Nothing in UK healthcare is credibly operating here, and no current or announced framework would support it.

Three horizons, and an honest caveat

Nought to twelve months. Three documents land, not one. The National Commission reports. The Medical Devices (Amendment) Regulations are adopted. And the Information Commissioner's draft guidance on automated decision-making goes out for consultation, ahead of a statutory code of practice on artificial intelligence.

That third one is the framework nobody in health is watching. SI 2026/425 came into force on **12 May 2026**, requiring the Commissioner to prepare a code of practice on "developing and using artificial intelligence" and on "automated decision-making". The ICO is separately engaging eleven foundation model developers on the lawfulness of their training data. Whatever the code says will reach ambient scribing, NHS App triage and every diagnostic tool in the country, because all of them process personal data. It is not a health framework and it will not settle the device-status question — but it is the only UK AI code with a statutory duty behind it, and an organisation preparing a position for the National Commission and stopping there will have prepared half of one.

One to three years. Post-market surveillance for adaptive systems moves from a gap everyone named to an obligation someone enforces. Predetermined change control plans are the likely mechanism. Organisations with their own instrumentation will find this cheap; the rest start from nothing.

Three years and beyond. Either liability is redistributed across developer, deployer and clinician, or it stays where NHS England put it in June 2025. Between 77% and 88% of respondents to the National Commission said the current position is untenable. That pressure resolves one way or the other.

The first horizon is a matter of public record. The second is a reasoned reading of what the Commission's own findings point towards. The third is a projection. Anyone presenting a three-year view of AI regulation as a forecast is selling something.

The accountability ladder is the author's framework. Regulatory anchors: MHRA guidance, 29 July 2026; UK GDPR Articles 22A–22D, in force 5 February 2026; SI 2026/425, in force 12 May 2026; National Commission findings, 11 June 2026.

ELEVEN

Benefit realisation

The benefit case for this technology currently has one side of the ledger filled in. Completing it costs almost nothing and cannot be done retrospectively.

A worked example

NHS England's trial reports documentation time down 51.7% and emergency department capacity up 13.4%. Applied to a five-hundred-clinician trust at a deliberately conservative one hour saved per clinician per week, that is roughly 26,000 clinical hours a year. The figure is real, it is defensible, and it is what a business case will be built on.

It is also incomplete, because the same trial measured no error rate. A benefit case that quantifies the upside precisely and leaves the downside unquantified is not a benefit case; it is a forecast with one variable missing. The four measures below fill the gap.

Correction rate	The proportion of AI-generated notes materially edited before sign-off. This is the closest available proxy for how often the system is wrong, and it is collected as a by-product of normal work.
Correction decay	Whether that rate falls over time when product accuracy has not improved. A falling correction rate against a static product is automation bias, measured.
Subgroup variance	Correction rate broken down by accent, first language and speech difference. This is where the equity risk lives and where the regulator says the problem appears only after deployment.
Attributable incidents	Incidents in which AI-generated documentation was a contributing factor, recorded as such locally because no national route does it.

WHY THE SECOND SET MATTERS MORE

Time saved is easy to measure, which is why it is the number every organisation has. Correction rate is the number that tells a board whether the safety case is holding, and it is the one that will be asked for the first time something goes wrong. It costs almost nothing to collect from the day of deployment and it cannot be reconstructed afterwards. An organisation that starts collecting it today has a two-year record by the time anyone requires one.

GOSH DRIVE v1.3, February 2026 for the trial figures. The measurement framework is the author's, applied across this briefing series.

TWELVE

A way in

Eight workstreams. The first three form a natural first step.

01	**Readiness assessment**	Where AI already is in the organisation, sanctioned and unsanctioned, and what evidence exists for each deployment	4–6 weeks
02	**Use case identification and triage**	Which uses are assistive, which are augmentive, and which have quietly become medical devices through how they are used	3–4 weeks
03	**Regulatory and legal position paper**	A defensible written position on device status, accountability and Articles 22A–22D, ready for both the Commission's recommendations and the ICO code	3–4 weeks
04	**Board and executive governance**	Who decides, who is accountable for an automated output, and what the board sees monthly	4 weeks
05	**Threat and integrity response**	Local incident recording for AI involvement, and a position on patients arriving pre-triaged by unregulated tools	4 weeks
06	**Benefit realisation framework**	The four measures in section eleven, instrumented from day one	3 weeks, parallel
07	**Build, buy or partner evaluation**	Supplier assurance that survives the device-status change, including what happens to products bought under the old rule	4–6 weeks
08	**Twelve-month mobilisation roadmap**	Sequenced, costed, and written to be handed to whoever runs it	2 weeks

THE FIRST STEP

The first three workstreams form a natural first step of six to eight weeks.

Two frameworks are coming. The National Commission's recommendations were due in summer 2026 and have not been published, and the ICO's code of practice on AI is being written under a statutory duty in force since May 2026. Six to eight weeks fits inside the time still available — and it is the last window in which an organisation forms its own position rather than receives one.

Contact: nc@neilcatton.com · neilcatton.com · writing.neilcatton.com

THIRTEEN

Sources

Every source read directly. Where a document could not be read in full, that is stated.

Regulators and government

- MHRA, *AI Airlock Sandbox Phase 2 Programme Report*, June 2026.
- MHRA, guidance on the regulatory status of ambient voice technologies, 29 July 2026.
- MHRA, *National Commission into the Regulation of AI in Healthcare: call for evidence findings*, 11 June 2026.
- MHRA, *Impact of AI on the regulation of medical products*, 30 April 2024, last updated 31 July 2026.
- MHRA, AI Airlock Phase 2 tool list, 16 October 2025.
- NHS England, priority notification on ambient voice, Alec Price-Forbes, National CCIO, 9 June 2025.
- NHS England, *Guidance on the use of AI-enabled ambient scribing products*, 27 April 2025, v3, updated 29 July 2026.
- NHS England, *Medical device regulation for ambient voice technology*, 29 July 2026.
- NHS England, ambient voice supplier registry, 19 January 2026.
- NHS England, "NHS accelerates AI rollout", 4 July 2026.
- NHS England, AI-based skin lesion analysis technology (DERM) implementation guidance.
- NHS England, Federated Data Platform FAQs.
- GOV.UK, *AI to speed up cancer diagnosis for millions of NHS patients*, 10 June 2026.
- NIHR Rapid Service Evaluation Team, rapid evaluation of AI in chest diagnostics, 16 March 2026.
- DSIT/GOV.UK, AI demand-forecasting tool for A&E, 28 December 2025.
- Office for Life Sciences, *Bioscience and health technology sector statistics 2023 to 2024*, 2 October 2025.
- ONS, *Artificial intelligence in UK businesses: 2023 to 2026*, 20 July 2026.
- National Commission into the Regulation of AI in Healthcare, gov.uk, updated 19 June 2026.
- HIQA (Ireland), *National Guidance for the Responsible and Safe Use of Artificial Intelligence in Health and Social Care Services*.
- Information Commissioner's Office, *The Data Use and Access Act 2025 — what does it mean for organisations?*, page updated 19 June 2026, and *AI and biometrics strategy update*, March 2026.

Legislation

- Data (Use and Access) Act 2025, s.80, inserting UK GDPR Articles 22A–22D, in force 5 February 2026. SI 2026/425, the Code of Practice on Artificial Intelligence and Automated Decision-Making Regulations 2026, in force 12 May 2026. Medical Devices Regulations 2002 (as amended), and the draft 2026 Amendment Regulations — not read in full; no provision quoted.

NHS trusts and research

- GOSH DRIVE, *The use of Ambient Voice Technology with Generative AI in Multiple Clinical Settings Across the NHS*, v1.3, 11 February 2026. NHS England-sponsored; product TORTUS.
- TheHill at Oxford University Hospitals NHS FT, *Ambient Voice Technology in Acute Services — An Evaluation*, February 2026. Trust-commissioned and funded.
- Nuffield Trust and RCGP, *How are GPs using AI?*, December 2025. 2,108 GPs, 28% response rate, fieldwork 28 July – 20 August 2025.
- Derksen et al., "Use of AI scribes in UK primary care", *npj Digital Medicine*, 16 May 2026. 598 GPs; Barts Charity and Cancer Research UK.

- Clark, Duffy, Kim, MacAskill, May and McCann, *The use of AI in UK healthcare*, KCL, May 2026. Focaldata, 2,093 UK adults, 24–30 March 2026.
- Our Future Health, cohort profile, *International Journal of Epidemiology*, 15 October 2025.

Trade bodies and patient safety

- ABHI, response to the MHRA call for evidence, February 2026.
- Patient Safety Learning, *Ambient voice technology: the patient safety risks*, Mark Hughes, 25 August 2026.
- MHRA statement to *Pulse*, 6 August 2025, on Yellow Card reporting.

This edition adds primary sources for the national imaging rollout, the NIHR evaluation of AI in chest diagnostics, the DERM autonomous skin-cancer pathway, MHRA Airlock's non-documentation tool cohort, and the data infrastructure — the Federated Data Platform, the A&E demand-forecasting tool and Our Future Health — underneath them, reflected in section four.

FOURTEEN

About the author

Neil Catton is an independent technology strategist with more than thirty-nine years of technology leadership across twenty-plus sectors. He has held CTO and Director-level roles at BT, Wipro, Fujitsu, Sopra Steria and Capita, and works across a fractional portfolio spanning public safety, insurance, health technology and early-stage companies.

In health specifically, he was chief architect for a national health platform serving 400 million citizens across eleven medical cities and 6,500 clinics, and led product strategy work for Capita's health business. The questions in this briefing — what the evidence actually shows, who is accountable for an automated output, and how a board proves a safety case it was never given the data for — are the questions he has spent a career answering in regulated environments.

He is the author of four books, including the trilogy *The Next Evolution*, *The Cognitive Crucible* and *The Shadow System*, writing at the intersection of technology, ethics and human purpose. He writes at writing.neilcatton.com.

Contact: nc@neilcatton.com · neilcatton.com · writing.neilcatton.com

How this briefing was produced

Every figure is traced to a regulator, a trade body, an ombudsman, published legislation, named academic research, or an operator speaking on its own record. Where a figure could not be traced to one of those, it was excluded and named in section one rather than quietly dropped. Where a published figure carries no denominator, that is stated alongside it. Where two credible sources disagree, both are given and the disagreement is reported. Three findings in earlier drafts were corrected before publication. An argument that incidents were going unreported was replaced by the structural finding in section six. A proportion was corrected to its true base. And a description of section 80 of the Data (Use and Access) Act 2025 as having permitted automated decision-making was corrected: the prohibition still applies where a decision rests on special category data, and the statutory code of practice on AI that the Information Commissioner has been required to prepare since May 2026 had been omitted altogether.

NOTICE

1. This briefing is not legal, regulatory, financial or professional advice.

2. The position is stated as at 8 September 2026. The National Commission's recommendations and the Medical Devices (Amendment) Regulations 2026 were both outstanding at that date and may since have moved it.

3. Third-party sources are cited in good faith. Verify before relying on them.

4. Named organisations appear only from regulator, trust or published research findings, and are used illustratively. None has been engaged or approached in connection with this briefing.

5. No advisory or client relationship is created by receipt of this document.

6. Independence. The author holds no vendor affiliations, reseller agreements, referral arrangements, commission interests or equity positions in anything referenced in this briefing, and has received no payment from any organisation named in it.

7. © Neil Catton 2026. Provided for the recipient's internal use.