Health Data Security: A Rapid Systematic Review of Public Understanding, Perceptions, and Resources

February 2025

Authors and project team

Kohlrabi team
Katie Tiley – Review lead
Dr Fran Harkness- Delivery Lead

Additional analysis:
Yaning Wu
Aaron Koay

Understanding Patient Data team
Emma Morgan – Research & Evidence Manager

Executive Summary

Background

This project is commissioned by Understanding Patient Data (UPD), hosted at the NHS Confederation, and undertaken by Kohlrabi from November 2024 to May 2025. The project involves a desk review, which then informed deliberative dialogues and co-creation workshops to develop specifications for public-facing health data security resources/explainers. Preliminary scoping research showed many existing explainers are directed at organisations to support security legislation, and public-facing explainers are unengaging and tend to be specific to situations where data has been leaked, shared, misused, etc.

This rapid systematic review was the first stage of this project, surveying the peer-review, grey literature, and media landscapes to address the following Research Questions (RQ):

1. What does the public currently understand about health data security?
2. What communication styles are used to talk about complex technical processes which address public concerns?
3. What are existing health data security resources?

Therefore, we screened both research about the public understanding of health data security (RQ1) and public-facing communications about health data security (RQ2 & RQ3). The findings of this review will be useful for academic, public and private bodies to establish key knowledge gaps in the research landscape of public understanding of health data security and inform development of effective public-facing health data security explainers.

Key Findings

⇒ Over 5,000 records were screened across twelve literature and media sources.

Public understanding and perceptions of health data security (RQ1):

- Reviews of peer-reviewed and grey literature suggest that **little is known about the UK publics' understanding of health data security** concepts and issues.
- The UK public tend to be supportive of health data sharing for direct care, though there are common concerns about the exploitation of health data by commercial companies and the risk of data breaches.

Public-facing communications about health data security (RQ2 & RQ3):

- Public-facing information is **overwhelmingly text-based and has a technical reading level** which is inaccessible to non-university level audience, regardless of publication source (public and private bodies or news agencies).
- Public and private bodies with responsibility for maintaining the security of health data used for direct care and/or research generally employed a neutral or positive sentiment in relevant documents.
- Media sources generally presented information about health data security with a negative sentiment and typically focus on malicious external security breaches.

Recommendations

- Academic and public bodies should directly research public understandings of health data security, including a diverse range of participant demographic characteristics and perspectives.
- Public and private bodies should develop public-facing resources relating to health data security **in the context of direct care**, acknowledging the provision for resources focused on security of health data sharing/access for research.
- Public-facing resources should be made **more accessible** by both improving the readability of text-based resources and integrating multimedia explainers.
- Public-facing resources should ensure they use a **neutral tone**, finding the middle ground between public bodies' tendency to be reassuring and the media tendency to be alarming.

Evidence Gaps

- All peer-reviewed research and most grey literature examined public understanding of health data security in the context of data sharing for research, however, there is a paucity of evidence of the UK publics' understanding in the context of direct care.
- Of over 5,000 records screened across twelve literature and media sources, only a single video resource and a single audio (podcast) resource were identified.

⇒ **There is a critical gap for non-technical public-facing multimedia educational resources about health data security in the context of direct care.**

Next Steps

The next steps involved a series of in-person and online workshops. The first set of workshops included an introduction to key concepts in health data security and further

established a foundation of what the public understands about health data security in the context of direct care. The second set of workshops focused on co-creation, developing recommendations for resources and identifying specifics like language choice, resource type, and style. These recommendations and specifications were then tested further with health and care professionals and key stakeholders through individual interviews.

Table of Contents

Introduction

The increasing digitisation of modern health systems has created ever-larger volumes of accessible health data relating to individual patients and procedures. There are vast opportunities to use this data to have a positive impact on care outcomes and quality of life; however, the storage and use of this data presents an ever-growing security risk. Alongside the development of robust systems to ensure data safeguarding, public awareness and trust in data access processes and data agencies is necessary to facilitate the optimal use of this personal data.

What is health data security?

We consider security in relation to the processes and practices involved in the protection of patient data (such as personal details, medical records, and treatment details etc.) from unauthorised access, disclosure, alteration, or destruction, through either accidental or malicious breaches.

This is in line with the ICO definition of a personal data breach:
"Under the UK GDPR, a personal data breach means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data – whether due to accidental or deliberate causes or organisations failing to take appropriate action."

Box 1. Defining health data security, in line with the Information Commissioner's Office (ICO) definition of personal data breach (1).

Most people trust the NHS (used here to generally mean all public health services across the UK) to keep their health data secure, even amongst those who are less trusting of the NHS more broadly (7). However, many remain concerned about security threats such as cyber-attacks and there is little awareness or understanding of what is done to keep health data safe (8). Even in research exploring public attitudes towards other health data topics, such as data linkage, the importance of data security is frequently emphasised by public participants (9).

While some research and polling on UK publics' understanding and perceptions of health data security has been produced, syntheses of evidence on these topics are dated. Reviews of public perceptions of secondary uses of NHS data (2) and of public responses to data sharing for research purposes in global settings (3) contain data from no later than 2021, despite the policy and technological landscape surrounding health data security having rapidly evolved. The ICO report that in 2024 (4) the most reported reasons for health data breaches were:

1. Other non-cyber incident (23.27%)

2. Hardware/software misconfiguration (15.46%)
3. Unauthorised access (14.43%) and
4. Data emailed to incorrect recipient (13.84%)

Unauthorised access may relate to staff internally accessing data without genuine reason, a high-profile example of this being hospital staff trying to access the medical records of Kate Middleton, the Princess of Wales (5). Cyber-attacks against NHS services, while not occurring as often as accidental breaches, are often widely publicised and may have greater impacts on the functioning of health systems, such as the attack on the pathology laboratory, Synnovis, in June 2024 (3).

NHS communications about data security, which understandably tend to promote positivity and reassurance, can be perceived particularly by those who are more 'disengaged and health data protective' as too emotive or pressuring (7), while media stories which tend to focus on breaches and cyber-attacks in an alarmist way (10) can spark panic and misunderstandings of risks and consequences (11). Between these two approaches of reassure and alarm is one that instead seeks to help people to understand the basic facts of health data security, to support them to make more informed choices about their own data, and to take a more critical approach to information given by various organisations.

Review Objectives

Therefore, we conducted a review of evidence from both peer-reviewed and grey literature resources with the following objectives:

1. **What does the public currently understand about health data security?** Highlight what the public currently understands about health data security (including issues such as cyber-attacks, breaches, and data handling) and where significant knowledge gaps or misconceptions may exist (RQ1);
2. **What communication styles are used to talk about complex technical processes which address public concerns?** Identify communication styles for how complex technical processes related to health data security (such as cybersecurity measures, data breaches, and the NHS' response to these threats) are communicated, including clarity, accessibility, tone, and how these resources address public concerns (RQ2);
3. **What are existing health data security resources?** Synthesise the volume of resources and communication, and their range across topics and sources, including the media (RQ3).

Methodology

Evidence searches were conducted across **twelve sources**. The **PubMed** database was searched for peer-reviewed literature on public understanding and/or perceptions of health data security (RQ1) (see Appendix A for search terms). Websites of public bodies working within the realm of health data security, including **NHS England**, **NHS Digital (**prior to its merge with NHS England in 2023 (12)), **GOV.UK**, and **Health and Social Care Committee**, were searched to identify grey literature on public understanding and/or perceptions (RQ1) and public-facing resources relating to health data security (RQ2 and RQ3) (Appendix B). To supplement evidence identified from these searches, UPD additionally made a **Call for Evidence** through its social media channels and networks on 25th November 2024 for resources relevant to all research questions. Finally, six news outlets, prioritised by readership number and balanced for political leaning, were searched for articles relating to health data security: **the BBC**, **the Daily Mail**, **the Guardian**, **the Telegraph**, **the Mirror**, and **Metro** (Appendix C). A second reviewer checked 20% of excluded resources and data extraction for quality assurance and to ensure consistency.

The following **inclusion / exclusion criteria** were used to select relevant records:

- publication date after 1st January 2019, chosen to prioritise most relevant records and ensure a manageable amount of results;
- published in English, relating to the UK context;
- provide sufficient detail to extract key findings (e.g. conference abstracts excluded);
- sample population comprises the general public including patients (e.g. a sample population of health care professionals excluded); and
- the record contains reference to health data security as part of its main findings.

For included research records investigating public understanding and/or perceptions of health data security (**RQ1**) we recorded:

- population characteristics (age, gender, ethnicity, disability status, sample description, sample size)
- context for investigating public understanding and/or perception
- method of assessing of public understanding and/or perception
- health data security domains assessed (short key words used for thematic analysis, e.g., "Confidentiality", "Data breaches")
- key findings.

For included public-facing health data security resources, we synthesised content, communication style and sentiment (**RQ2 and RQ3**) and recorded:

- the context for providing information;
- content type (e.g., text, audio, visual);
- health data security domains;

- content sentiment (categorical scale: Positive, Neutral-leaning-positive, Neutral, Neutral-leaning-negative, Negative);
- resource type (e.g., blog, news, press release, etc.); and
- reading level assessed using the Flesch Reading Ease Score (9,10).

Findings

Search and Screening

The search of PubMed identified 323 peer-reviewed literature sources which were screened, of which 12 met the inclusion criteria. Across all grey literature sources (NHS England, NHS Digital, GOV.UK, and Health and Social Care Committee), 1,042 records were screened, of which 73 met the inclusion criteria. A respondent to the UPD Call for Evidence provided three records of which two met the inclusion criteria. Finally, 3,789 records were identified and screened across the six news outlets, leading to 41 included media articles. A total of 5,157 records were screened, leading to 128 included records with complete data extraction.

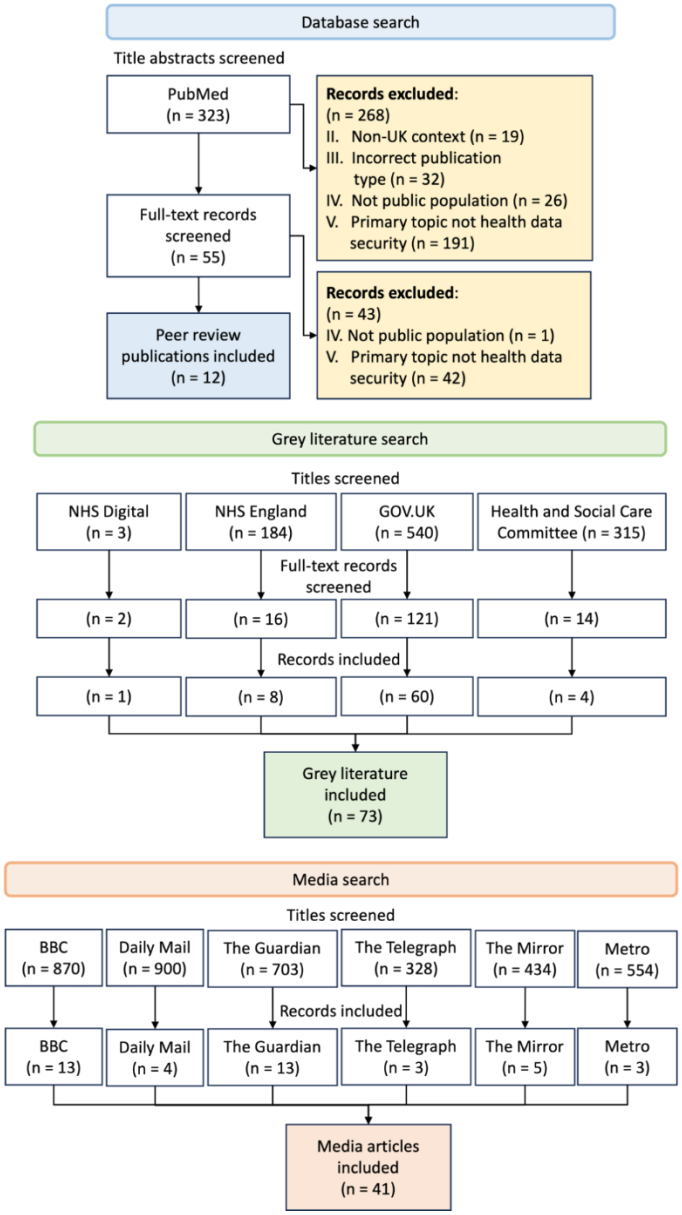

Figure 1. Flow diagram showing numbers of screened and included records from all sources.

RQ1: What does the public currently understand about health data security?

Our search identified 12 peer-reviewed studies and 18 grey literature publications which examined public understanding and/or perceptions of health data security issues.

i. Participants

Despite peer-reviewed publications seeking the views of the public, they generally included the views of people who had higher than average exposure to the health system, for example, either people with lived experience of seeking healthcare, caring for loved ones who had sought healthcare, or participating in health-related research. In comparison, grey literature resources evaluated views of the public with varying levels of engagement with the healthcare system and health research. Demographic characteristics of public participants, such as age, sex, ethnicity, and disability status, were rarely reported across peer-reviewed studies and those which did reported majority White adults with even gender balance. Grey literature resources surveyed participants who were assumed to be "representative of the British population", providing few details of demographic characteristics.

ii. Research Context

Though all included resources reported some findings relating to public understanding and/or perceptions of health data security, they were generally not the main aim of the study. Peer-reviewed publications reported these findings qualitatively and in a tangential manner alongside other research questions, such as public views of research participation (15) or of emerging mobile health technologies (16), and grey literature resources similarly reported findings within a broader context, such as data access or digitisation of health services. Only public surveys/polling provided descriptive quantitative statistics, for example, polling for public opinion on data sharing during the COVID-19 pandemic gives statistics such as "A majority (64%) said that they would trust government agencies to use information about them such as coronavirus test results... 17% did not agree with this and 19% were not sure" (12).

Most peer-reviewed and grey literature sources discussed public views surrounding security for types of health data that were routinely collected in the care setting, such as medical history, while a handful of publications reported on genomic data and health data extracted from mobile devices (two data types not employed in standard care in the UK). In addition, most peer-reviewed and grey literature sources focused on public perceptions or understanding in relation to hypothetical or future health data security issues instead of issues relating to current uses of participants' health data.

Health Data Security...
Understanding: "I do/don't know how my data is stored when it's being used for X"
vs
Perception: "I don't want X party to be doing X with my data or storing my data in X"

Box 2. Demonstrating nuanced differences between understanding and perception in health data security.

*iii. Findings: public **understanding** of health data security*

A total of eight records reported findings relating to public understanding of health data security; across peer-reviewed publications (17–21) and grey literature resources (2,22,23). While the peer-reviewed publications provide tangential evidence of some public awareness of relevant concepts, the grey literature suggests that greater public understanding is required, particularly in relation to health data security in direct care as all the evidence relates to health data security in relation to research. For example, peer-reviewed publications found participants were aware of issues surrounding open access health data, disease-specific patient registries, confidentiality agreements, and a need to fill gaps in understanding, each topic closely related to their individual involvement with the health system. On the other hand, grey literature suggests that public understanding is insufficient around concepts such as aggregate vs individual data, anonymous vs pseudonymous data, uses of health data by commercial parties, and existing safeguarding guidelines and mechanisms. One resource reporting on polling during and before the COVID-19 pandemic also showed that public awareness of the National Data Opt-Out scheme fluctuated over time, with less than half of surveyed participants familiar with the scheme in July 2020 (23).

"Participants reported awareness that study data would be kept confidential ('I guess because all of the information is like private' (Male/17/High-Risk/Control)) and that confidentiality breaches within this sensitive context 'could affect that person's, say, chances of getting a job or something' (Male/17/High-Risk/Control)."

Box 3. Example of a study which asked adolescents in alcohol intervention trials about health data security. While this is the best quality reporting of public opinion in peer-reviewed publications, it asks about data sharing/access for research purposes during a clinical trial, not for direct care.

*iv. Findings: public **perceptions** of health data security*

A total of 26 records reported data relating to public perceptions of health data security across peer-reviewed and grey literature. Public perceptions were categorised into:

(1) substantial support of and trust in the secure sharing/access of health data for research and/or direct care use,
(2) conditional support for the sharing/access of health data given certain caveats relating to security and confidentiality, and
(3) concerns surrounding the sharing/access of health data given worries or fears surrounding the misuse of data and resulting negative consequences.

Among resources reporting substantial public support, participants found the sharing of health records, clinical samples, consultation data, and health behaviour data for research use generally acceptable, potentially perceiving little security risk due to rigorous data practices in research and in principle supporting the intent of researchers to use health data in the pursuit of wider public health benefits. Some grey literature resources in this category additionally mention public support for sharing health data specifically for healthcare practitioners' reflective practice - again, a case where the intent is to improve healthcare. Conversely, resources reporting public concerns cite core issues to be improper storage and potential misuse of genetic data, privacy breaches in the context of mHealth (mobile) data, and the granting of data access to private/commercial companies, cases where malicious intent is perceived as the greatest security risk either from individuals or companies and data is either perceived as more desirable (genetic data) or less secure (collection of large amounts of data via a mobile phone). The resources reporting conditional public support for sharing health data provide the most insight into what the public requires from responsible organisations:

- provision of safeguarding information;
- option to consent to sharing only specific types of health data;
- ability to share data in a confidential environment or context;
- trust in the organisation and/or individual using the data;
- control over which data will be used and how that data will be used;
- a clear public benefit derived from the applicable use of health data;
- clear and transparent (i.e., balanced) communications about the use of health data;
- restrictions on the use of health data by commercial companies;
- use of secure data environments as a preferred method of data sharing for research;
- regulation of secondary uses of health data to enhance cyber security;

- patient and public involvement (PPI) in determining whether those secondary uses could go ahead as proposed.

Furthermore, four resources noted public support for the introduction of the eighth principle of the Caldicott Principles (24) to inform service users more fully of the secondary uses of their health data.

iv. Research Gaps

The evidence largely ignores data sharing/access in the context of direct care and its potential for breaches, and instead overwhelmingly considers data security around research. The numerous categories of health data breach incidents reported to the ICO (4) highlight the diverse scenarios through which health data security can be compromised and which the public are largely unaware of, such as non-cyber incidents (e.g., incorrect disposal of paperwork) and internal accidents (e.g., data emailed to the incorrect recipient). UPD have previously published a report on how to communicate with the public about Secure Data Environments and Trusted Research Environments (25); although the findings in that report contain useful learnings for health data security for direct care, including the importance of discussing data security with the public and not assuming any prior technical knowledge in public communications, there are significant gaps in our understanding of public attitudes and knowledge in relation to data security in direct care specifically.

RQ2 & RQ3: Existing health data security resources / explainers and their communication styles

A total of 72 public-facing resources on health data security were identified. These documents originated primarily from public bodies, with only two resources originating from a social enterprise working with health data. Documents included blog posts, press releases, privacy notices, policy papers, government reports, speeches, and meeting minutes.

i. Medium, Audience, and Readability

All resources were entirely text-based, including three transcriptions of speeches by UK government figures, except for one video explainer on data transparency (26), and a resource containing helpful diagrams outlining implementation milestones of a government data strategy (27). All included resources were assumed to be for a public audience or at least in the public interest by virtue of their public availability and explicit indication of a target audience.

Figure 2. Screenshot of an example text-based explainer from the UK Health Security Agency (UKHSA) (28)

The reading level of included resources, i.e., ease of readability with respect to word complexity and sentence length, varied by source and content. However, with few exceptions, resources scored between 20-50 out of 100, implying a reading level compatible with undergraduate degree materials (lower scores indicate a more difficult or technical read). Notably, three speech transcriptions achieved Flesch Reading Ease scores (13,14) of 50-60 out of 100, suggesting a reading level compatible with GCSE or A-level education. In addition, materials that were arguably more likely to be accessed by the public (e.g., blog posts, press releases, and news stories) did not have reading levels more accessible to a wider audience than technical materials, such as policy papers and reports. Data from 2023 showed that 84% of UK-based adults hold GCSEs (29), meaning that a reading score of 50+ may be accessible to the majority of UK audiences.

ii. Content

The public-facing resources which met the inclusion criteria covered a variety of health data security domains. These included:

- opt-out systems for data sharing (predominantly for research purposes);
- storage and pseudonymisation of health data;
- trusted research environments and secure data environments;
- information governance;
- cyber security training for healthcare and healthcare-adjacent staff.

A substantial proportion of resources published from 2020-2022 focused on data sharing within the context of COVID-19-related care and/or research. Eleven resources contained information relating to the impact of, prevention of, and recovery from data

breaches or cyber-attacks affecting healthcare data in the UK. These resources focused predominantly on malicious external attacks, possibly reflecting the public's greater awareness of high-profile and high-impact malicious data breaches (discussed more in the media analysis), and is misaligned with ICO data showing some of the most common health data breaches to be accidental and internal, such as data being emailed to incorrect recipients (30). It was common for resources to reference relevant legislation such as the General Data Protection Regulation (GDPR).

iii. Sentiment

Resources' sentiment was assessed as the tone or attitude of its author(s) towards the referenced domains of health data security. There was a notable dichotomy between sources; for example, press releases, news stories, and reports were predominantly positive or neutral-to-positive-leaning from NHS England, NHS Digital, or GOV.UK sources (including sources from the Department of Health and Social Care, the Information Commissioner's Office, the National Data Guardian, Public Health England, and the UK Health Security Agency). In contrast, resources from the Health and Social Care Committee, which commissions an independent expert panel to assess "policy, spending, and administration" within the UK's Department of Health and Social Care, were characterised as neutral or neutral-to-negative-leaning sentiment. Privacy notices and similar documents describing the processing of public users' data were uniformly neutral in sentiment.

RQ2 & RQ3: Media articles and their communication styles

A total of 41 media resources relating to health data security were identified, with most included articles published by the BBC and the Guardian (13 articles each).

i. Medium, Readability, and Audience

All resources except for one podcast were textual (31). Most media resources had a Flesch Reading Ease score between 40-50 out of 100 (range 29.5 – 67.6), suggesting these articles used only slightly more accessible language than health data security resources from the websites of public bodies.

ii. Content

Most news articles related to health data security breaches, with the majority reporting on breaches by malicious external, rather than internal, actors. Impacts on patients and organisations described in these articles included:

- longer wait times in primary, secondary, and emergency care;
- cancellations and/or delays in appointments, blood tests, and operations;
- financial loss for both organisations and patients (with some patients opting for tests/procedures in private centres due to exorbitant wait times); and
- stolen patient data, with subsequent impacts on clinical decision-making.

However, less than half of articles provided direct instructions to individuals affected by security breaches and any available instructions were typically non-specific. For example, being generally "on guard" for future breaches and sharing views about the future of NHS health data security on social media channels or surveys run by public healthcare bodies. Specific instructions encouraged the public to use official channels (such as NHS websites) to find verified information about care providers, bring paper documentation to appointments rather than rely on electronic records, refrain from using emergency services in the absence of emergency conditions, or in one notable case, contact specific incident helplines with queries (32).

Three more hospitals hit by cyber attack

GOOGLE

Alder Hey Children's Hospital, Liverpool Heart and Chest Hospital and Royal Liverpool University Hospital were affected by the attack on Thursday

Angela Ferguson & Lynette Horsburgh
BBC News, Merseyside

4 December 2024

Figure 3. Screenshot of an example news article about a cyber-attack from the BBC in 2024 (33)

Where no breaches were mentioned, articles described either concerns over NHS data security by public bodies or plans to improve that security. Aside from one article discussing health data security for research (34), all other resources described health data security issues in relation to data use for direct care and it was rare for them to mention health data security legislation. This notably contrasts the peer reviewed and grey literature research that overly focuses on health data security for research. While the public can access some information about health data security for direct care

through government reports, and less commonly, media articles, we don't know how much these resources are accessed and consumed by the public or what readers may take away from them.

Starmer insists NHS must make better use of tech amid data protection concerns

By PA MEDIA
PUBLISHED: 11:57, 21 October 2024 | **UPDATED:** 11:57, 21 October 2024

Share

The NHS must make "much more use of technology", the Prime Minister has insisted amid concerns about plans to share patient data in the NHS.

New laws are set to be introduced to make patient records available across all NHS hospitals, GP surgeries and ambulance services in England.

And plans for a "single patient record" have been unveiled, which will summarise all of a patient's health information, test results and letters in the NHS App, the Department of Health and Social Care said.

Figure 4. Screenshot of an example news article about new health data security policy from the Daily Mail in 2024 (35)

iii. Sentiment

Most media sources demonstrated neutral and neutral-to-negative-leaning sentiment, with one article expressing exclusively negative sentiment: "Beijing 'trying to harvest NHS health data from British patients to develop bioweapons', experts warn" (36). Articles with a neutral tone overwhelmingly reported guidance from public bodies or public servants, whereas articles with negative sentiments were characterised by quotes from public servants or the public on the impact, or potential impact, of attacks on individuals and systems.

Conclusion

This rapid systematic review demonstrates that public understanding of health data security is rarely elicited by published research. The limited evidence shows generally poor public understanding; however, individuals consider confidentiality and privacy important issues when deciding whether data should be shared and have concerns surrounding the use of health data by commercial companies and the potential for data breaches.

Publicly available resources on health data security are almost exclusively text-based and have a technical reading level which is inaccessible to non-university level audience, regardless of publication source (public and private bodies or news agencies). Resources from public and private bodies largely focus on health data security in the context of data sharing for research and generally embody a positive sentiment, whereas media articles focus on health data security in the context of direct care and embody negative-leaning sentiment, likely due to their predominant focus on data breaches from malicious external actors.

These findings largely support our initial perception that NHS communications tend to promote positivity and reassurance, while media stories focus on security breaches and cyber-attacks in an alarmist way. In addition, this review shows there is a middle-ground between these sentiments found in resources from all sources which report on health data security guidance and legislation. However, despite being publicly available, these neutral explainers are consistently inaccessible to the public due to technical reading level and content density. This evidence confirms a dire need for public explainers which help people understand the basic facts of health data security particularly in relation to direct care, to support them to make more informed choices about their own data, and to take a more critical approach to information given by various organisations.

Key Take-Aways

- Most research into health data security is related to sharing / access for research, not direct care.
- The limited evidence available suggests the public know very little about health data security for direct care.
- Publicly available information is overwhelmingly text-based and technically inaccessible.
- Grey literature and media articles mostly focus on malicious external data breaches, despite the ICO reporting internal breaches are more common.

Recommendations

<u>Recommendation 1:</u> Gather more data on public understanding of health data security concepts among diverse demographics

- Most studies and surveys identified in this review only tangentially reported on public understanding of health data security.
- One peer-reviewed study with substantial focus on public understanding included a highly selected population (adolescents in alcohol intervention trials).
- More efforts are needed to collect quantitative and qualitative data on public knowledge of health data security concepts, policies, and actors.
- This data should be collected from nationally representative samples, with scope for oversampling of demographic groups currently underserved by health research and/or systems.

<u>Recommendation 2:</u> Develop public-facing resources on the security of health data use in the context of direct care

- This review identified a dearth of health data security resources relating to direct care, most focused on data security in relation to data sharing for secondary purposes, namely research.
- Developing more explainers on health data security in the context of direct care may improve public awareness and engagement in health systems. Despite potential perceptions that this type of data sharing/access is more implicitly acceptable given "no data leaves the system", this is not always the case as evidenced by the ICO incident reports for accidental breaches and malicious cyber-attacks.

<u>Recommendation 3:</u> Improve text accessibility of public health data security resources

- Public-facing documents relevant to health data security from all sources (public and private bodies and news agencies) had relatively advanced reading levels, potentially rendering them too technical to be truly accessible for the UK public.
- Readability for text-based health data security explainers for the public could be improved or offering easy-read versions could be widely adopted, developed by and for the public.

<u>Recommendation 4:</u> Develop multimedia public-facing health data security resources

- Almost all public-facing documents relevant to health data security across all sources (public and private bodies and news agencies) were text-based.

- Widely available and easily searchable multimedia resources, such as those produced by reputable public bodies, may raise public awareness of health data security issues.
- In all cases, diverse media should be rendered accessible to a wide variety of publics through alternative formats, such as videos, animations, and podcasts.

Acknowledgements

Kohlrabi and UPD would like to acknowledge the commitment of the steering group participants in their advisory position during the entire project process. Please note, their contribution is not an endorsement of the outputs of this project. We also thank the respondents to the Call for Evidence.

Project steering group members

Phil Huggins, Director for Cyber Policy & National Chief Information Security Officer for Health & Care -Department of Health & Social Care
Rachel Clarke, Senior Policy Officer - Information Commissioner's Office
Dr Nicola Byrne/ Ryan Avison, National Data Guardian/ NDG Head of Office
Dr Alison Knight, Data and Privacy Specialist – Health Research Authority
Dr Frances Burns, Programme Lead - HSC Northern Ireland Trusted Research Environment, Co-Founder and Director– Northern Ireland Public Data Panel
Alex Newberry, Head of Research Involvement, Governance & Informatics - Health & Care Research Wales, Welsh Government
Mark Simpson, Policy Team Leader, Inclusion and Engagement - Digital Health & Care Scotland, Scottish Government
Dr Mahi Hardalupas, Senior Policy Advisor (Data) - The Royal Society
Dr Amir Mehrkar, NHS GP, Senior Clinical Researcher - University of Oxford, Director of Information Governance and External Relations - Bennett Institute for Applied Data Science
Pritesh Mistry, Fellow (Digital Technologies) - The King's Fund
Richard Ballerand, Executive Group Member - UseMyData

We would also like to acknowledge contributions from the remaining experts and stakeholders who chose to remain anonymous.

References

1. ICO. What counts as a personal data breach? [Internet]. 2024 Apr [cited 2025 Jan 30]. Available from: https://ico.org.uk/media/about-the-ico/disclosure-log/4030826/c-297350-g8j0-knowledge-base-articles.pdf
2. NHS and the Patient Experience Library. Public perceptions of NHS data use [Internet]. 2021. Available from: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/09/bupa-fined-175-000-
3. Aitken M, De St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Med Ethics. 2016;17(1):1–24.
4. ICO. Information Commissioner's Office. 2024 [cited 2025 Jan 30]. Data Security Incident Trends - Proportion of Incidents Reported Dashboard. Available from: https://ico.org.uk/action-weve-taken/data-security-incident-trends/
5. The Independent. Hospital staff where Kate Middleton had surgery 'tried to access her medical records.' The Independent [Internet]. 2024 Mar 20 [cited 2025 Jan 30]; Available from: https://www.independent.co.uk/news/uk/home-news/kate-middleton-surgery-hospital-security-breach-b2515345.html
6. NHS England. Synnovis cyber attack – statement from NHS England [Internet]. 2024 Jun [cited 2024 Jan 19]. Available from: 19/01/2025https://www.england.nhs.uk/2024/06/synnovis-cyber-attack-statement-from-nhs-england/
7. NHS. Public attitudes to data in the NHS and social care [Internet]. 2024 May [cited 2025 Jan 19]. Available from: https://digital.nhs.uk/data-and-information/keeping-data-safe-and-benefitting-the-public/public-attitudes-to-data-in-the-nhs-and-social-care
8. Understanding Patient Data. Understanding public expectations of the use of health and care data [Internet]. 2019 [cited 2025 Jan 19]. Available from: https://understandingpatientdata.org.uk/sites/default/files/2019-07/Understanding%20public%20expectations%20of%20the%20use%20of%20health%20and%20care%20data.pdf
9. Jones LA, Nelder JR, Fryer JM, Alsop PH, Geary MR, Prince M, et al. Public opinion on sharing data from health services for clinical and research purposes without explicit consent: An anonymous online survey in the UK. BMJ Open. 2022;12(4).
10. Understanding Patient Data. Analysis of UK reporting on health data [Internet]. 2021. Available from: www.portland-communications.com
11. BBC. Cyber expert urges against "panic" over NHS data leak. BBC News [Internet]. 2024 May 8 [cited 2025 Jan 19]; Available from: https://www.bbc.co.uk/news/articles/clw08q19n9ro
12. NHS Digital. Health Education England, NHS Digital and NHS England merger [Internet]. 2023 Mar [cited 2025 Jan 30]. Available from: https://webarchive.nationalarchives.gov.uk/ukgwa/20231219181848/https:/digital.nhs.uk/about-nhs-digital/nhs-digital-merger-with-nhs-england
13. Flesch R, Ferry D. A New Readability Yardstick. Journal of Applied Psychology. 1948;32(3).

14. Microsoft. Microsoft Support. 2024 [cited 2025 Jan 10]. About Flesch Kincaid Readability and level statistics. Available from: https://support.microsoft.com/en-gb/office/about-flesch-kincaid-readability-and-level-statistics-8fbb787e-4e1d-4dcc-933f-c7d2f06e76ac
15. Fylan B, Munro J, O'Hara JK, Khatoon B, Lawton R. Developing a research community within an online healthcare feedback platform. Health Expectations. 2023;26(2):705–14.
16. Simblett SK, Bruno E, Siddi S, Matcham F, Giuliano L, López JH, et al. Patient perspectives on the acceptability of mHealth technology for remote measurement and management of epilepsy: A qualitative analysis. Epilepsy and Behavior. 2019;97:123–9.
17. Lynch E, McGovern R, Elzerbi C, Breckons M, Deluca P, Drummond C, et al. Adolescent perspectives about their participation in alcohol intervention research in emergency care: A qualitative exploration using ethical principles as an analytical framework. PLoS One. 2019;14(6).
18. Monticelli M, Francisco R, Brasil S, Marques-da-Silva D, Rijoff T, Pascoal C, et al. Stakeholders' views on drug development: the congenital disorders of glycosylation community perspective. Orphanet J Rare Dis. 2022;17(1).
19. Sim YJ, Townsend RF, Mills S, Stocker R, Stevenson E, McEvoy C, et al. Understanding engagement in diet and dementia prevention research among British South Asians: a short report of findings from a patient and public involvement group. Journal of Human Nutrition and Dietetics. 2024;37(4):899–908.
20. Reynolds J, Beresford R. "An Active, Productive Life": Narratives of, and Through, Participation in Public and Patient Involvement in Health Research. Qual Health Res. 2020;30(14):2265–77.
21. Minogue V, Cooke M, Donskoy AL, Vicary P. The legal, governance and ethical implications of involving service users and carers in research. Int J Health Care Qual Assur. 2019;32(5):818–31.
22. National Data Guardian. National Data Guardian 2022-2023 report [Internet]. 2024 Jan [cited 2025 Jan 19]. Available from: https://www.gov.uk/government/publications/national-data-guardian-2022-2023-report/national-data-guardian-2022-2023-report
23. National Data Guardian. Polling indicates growing public understanding about importance of using health and care data [Internet]. 2020 Oct [cited 2025 Jan 19]. Available from: https://www.gov.uk/government/news/polling-indicates-growing-public-understanding-about-importance-of-using-health-and-care-data
24. National Data Guardian. The Eight Caldicott Principles [Internet]. 2020 Dec [cited 2025 Jan 10]. Available from: https://www.gov.uk/government/publications/the-caldicott-principles
25. Understanding Patient Data. Understanding Patient Data: "What words to use when talking about health data" Rapid Evidence Review [Internet]. 2024 [cited 2025 Feb 6]. Available from: https://understandingpatientdata.org.uk/sites/default/files/2024-02/What%20Words%20To%20Use%20-%20Rapid%20Review.pdf
26. Optimum Patient Care Research Database. Optimum Patient Care Research Database. 2024 [cited 2024 Dec 5]. OPCRD Data Transparency. Available from: https://opcrd.optimumpatientcare.org/data-transparency
27. Dept of Health and Social Care. Care data matters: a roadmap for better adult social care data [Internet]. 2023 Dec [cited 2024 Dec 2]. Available from:

https://www.gov.uk/government/publications/care-data-matters-a-roadmap-for-better-adult-social-care-data/care-data-matters-a-roadmap-for-better-adult-social-care-data
28. UKHSA. UKHSA data strategy [Internet]. 2023 Sep [cited 2025 Jan 30]. Available from: https://www.gov.uk/government/publications/ukhsa-data-strategy/ukhsa-data-strategy
29. Dept for Education. Education and training statistics for the UK [Internet]. 2024 Nov [cited 2025 Jan 12]. Available from: https://explore-education-statistics.service.gov.uk/find-statistics/education-and-training-statistics-for-the-uk
30. National Data Guardian. National Data Guardian 2023-2024 report [Internet]. 2024 Dec [cited 2025 Jan 20]. Available from: https://www.gov.uk/government/publications/national-data-guardian-2023-2024-report/national-data-guardian-2023-2024-report
31. Radio4. Inside Health. 2024 [cited 2024 Dec 10]. Cancer vaccine trials and planning for cyber attacks. Available from: https://www.bbc.co.uk/programmes/m0022c2z
32. The Guardian. What does the London NHS hospitals data theft mean for patients? Cybercrime [Internet]. 2024 Jun 21 [cited 2025 Jan 20]; Available from: https://www.theguardian.com/technology/article/2024/jun/21/what-does-the-london-nhs-hospitals-data-theft-mean-for-patients
33. BBC. Three more hospitals hit by cyber attack. BBC [Internet]. 2024 Dec 4 [cited 2025 Jan 30]; Available from: https://www.bbc.co.uk/news/articles/c3vrk2e0xvwo
34. BBC. Improved privacy for children's wristband scheme. BBC News, Guernsey [Internet]. 2024 Oct 8 [cited 2025 Jan 20]; Available from: https://www.bbc.co.uk/news/articles/ced0wqjgq9eo
35. Daily Mail. Starmer insists NHS must make better use of tech amid data protection concerns. Daily Mail [Internet]. 2024 Oct 21 [cited 2025 Jan 30]; Available from: https://www.dailymail.co.uk/wires/pa/article-13982991/Starmer-insists-NHS-make-better-use-tech-amid-data-protection-concerns.html
36. The Mail. Beijing "trying to harvest NHS health data from British patients to develop bioweapons", experts warn. Mail Online [Internet]. 2024 Dec 22 [cited 2025 Jan 20]; Available from: https://www.dailymail.co.uk/news/article-14217957/Beijing-harvest-NHS-health-data-British-patients-bioweapons.html
37. Semrush. Open.Trends. 2024 [cited 2024 Dec 23]. Top websites in the United Kingdom (Newspapers Industry). Available from: https://www.semrush.com/trending-websites/gb/newspapers
38. YouGov. YouGov. 2017 [cited 2024 Dec 23]. How left or right-wing are the UK's newspapers? Available from: https://yougov.co.uk/politics/articles/17715-how-left-or-right-wing-are-uks-newspapers
39. PressGazette. Most Popular Newspapers UK ABC Monthly Circulation Figures. PressGazette [Internet]. 2025 Dec 23 [cited 2024 Dec 23]; Available from: https://pressgazette.co.uk/media-audience-and-business-data/media_metrics/most-popular-newspapers-uk-abc-monthly-circulation-figures-2/

Appendix A

Peer-review Database Search

We searched PubMed from 1st January 2019 to 4th December 2024 for English-language UK-based primary peer-reviewed studies of public understandings of health data security (RQ1) using a search strategy detailed in Table A.

Table A: Search strategy for peer-reviewed literature (RQ1)

Research question component	Search terms
Research relating to the public	(Patient Participation[MeSH Terms] OR Community Participation[MeSH Terms] OR lay[tiab] OR plain English[tiab] OR plain language[tiab] OR nontechnical description*[tiab] OR non-technical description*[tiab] OR public member[tiab])
Research relating to health data security	((“Computer Security”[MeSH Terms] OR data security[tiab] OR information security[tiab] OR cybersecurity[tiab] OR data protection[tiab] OR safeguard*[tiab] OR data sharing[tiab] OR data acquisition[tiab] OR data access[tiab] OR data safeguard*[tiab] OR data breach[tiab] OR data leak[tiab] OR compromised data[tiab] OR trusted research environment*[tiab] OR secure data environment*[tiab] OR confidenti*[tiab] OR priva*[tiab] OR anonymity[tiab] OR data regulation[tiab] OR record*[tiab] OR research*) AND (health*[tiab] OR disease[tiab] OR illness[tiab] OR hospital*[tiab] OR health service*[tiab] OR surger*[tiab]) OR (health data[tiab]))
Research relating to knowledge and/or understanding	(understand*[tiab] OR know*[tiab] OR literac*[tiab] OR comprehension[tiab])
Research based in the UK context	(United Kingdom[MeSH Terms] OR England[MeSH Terms] OR Wales[MeSH Terms] OR Northern Ireland[MeSH Terms] OR Scotland[MeSH Terms] OR Great Britain[MeSH Terms] OR National Health Service[MeSH Terms] OR United Kingdom[tiab] OR England[tiab] OR Wales[tiab] OR Scotland[tiab] OR Northern Ireland[tiab] OR London[tiab] OR NHS[tiab] OR National Health Service[tiab])

Appendix B

Grey Literature Search

We used the search term “health data security” across NHS Digital, NHS England, GOV.UK, and Health and Social Care Committee sources. Given the breadth of GOV.UK publications, we used additional search terms of "trusted research environment", "secure data environment", "data breach", “confidentiality”, and "data security" for its associated sources.

Table B: Web sources searched for grey literature (RQ1, RQ2)

Name of source	URL
NHS Digital*	https://digital.nhs.uk/
NHS England	https://www.england.nhs.uk/
GOV.UK**	https://www.gov.uk/
Health and Social Care Committee***	https://committees.parliament.uk/committee/81/health-and-social-care-committee/

* Includes documents categorised under “Service users and the public” within the “Audience” filter

** Includes documents categorised under “Guidance and regulation”, “News and communications”, “Research and statistics”, “Policy papers and consultations”, and “Transparency and freedom of information releases” from the Department of Health and Social Care, the Information Commissioner’s Office, the National Data Guardian, Public Health England (1 January 2019 until 30 September 2021 only), and the UK Health Security Agency.

*** Includes publications categorised under “Reports, special reports, and government responses” and “News”, with all available publications for the latter screened without searching due to the lack of search functionality on website

Appendix C

Media Search

To fully address RQ2 and RQ3, we searched for online articles across six media and news outlets. The search terms “NHS”, “cyber security”, and “data privacy” were used across each source; due to the abundance of records, the search was restricted to 2024. Considering the digital nature of news consumption, five news outlets were selected based on highest online readership in November 2024 (37) and balanced for political leaning, informed by YouGov (38): the BBC (bbc.co.uk), the Daily Mail (dailymail.co.uk), the Guardian (theguardian.com), the Telegraph (telegraph.co.uk), the Mirror (mirror.co.uk). To be inclusive of less digitally confident demographics, Metro (metro.co.uk) was additionally included as it has the highest physical print circulation in November 2024 (39). For each news outlet, articles were sorted by relevance and the first 300 articles screened.