

28 January 2022

Submission to UK Parliament Inquiry and Call for Evidence

The right to privacy: digital data

Key points

- The world is on the cusp of a data revolution in biomedical research and health and care design and delivery. We are already witnessing the value and transformative impact that data can have on our health and lives and the UK is well positioned to make a distinctive contribution to global health data science.
- Large-scale linkage and analysis of health-relevant data, across organisational and sovereign boundaries, is vital to deliver insights which can help address not only pandemic recovery but also the social and economic security of the global population.
- To fully realise the benefits of health data research at scale, we must create an ecosystem that includes all custodians of health-relevant data including government, industry, charities, universities and research institutions. This data ecosystem, if successfully convened and coordinated, will act as the foundation for a new era of health data science to advance the discovery and practice of medicine and care.
- Demonstrating trustworthiness in the safe, secure and innovative use of data is vital to achieve this ambition. The apparent high levels of public distrust warrant the need for secure regulatory standards and transparency over the uses and sharing of medical data. **This will require not only open, transparent public dialogue but also meaningful public involvement in the design and development of data initiatives.**
- HDR UK is committed to an approach to data access for research based primarily around the five safes framework.¹ This endorses the use of Trusted Research Environments (TREs)² which are secure computing environments that provide remote access to health data and support the highest standards of information governance, transparency and security by removing the need for data to be physically shared between different users. **Promotion of the five safes framework (see paragraph 12) will enable the right to privacy whilst unlocking the power of data.**
- Information governance processes are currently a significant barrier to wider access to and linkage of health data. Processes for accessing health data vary by custodian and are often slow and challenging to navigate. **Simplifying and streamlining these processes, including a single approach to the scrutiny of requests across multiple data controllers, has the potential to dramatically speed up data access and enable data linkage in a way that is transformational for the health data research community and patients.**
- **Clear, specific guidance from the Information Commissioners Office (ICO) stressing that the re-use of health data sets for research purposes is lawful, subject to relevant considerations and safeguards, would assist in building the confidence of data custodians in trustworthy data sharing.**

Introduction

1. Health Data Research UK (**HDR UK**) is the national institute for health data science, established by a consortium of nine UK funders³ in 2018. HDR UK's charitable mission is to unite the UK's health data to enable discoveries that improve people's lives. HDR UK's focus to date includes:
 - a) The assembly of a UK-wide data infrastructure for health research, including:
 - Bringing together over 60 of the UK's leading health and research organisations to collaborate on standards for data research – the UK Health Data Research Alliance⁴
 - Creating the UK's portal for the discovery and access to health data – the Health Data Research Innovation Gateway⁵
 - Establishing nine health data hubs based on partnerships across the NHS, academia and industry, to provide expert data research services
 - b) Accelerating the pace and scale of health and biomedical data science
 - Delivered 1,500 research papers with more than 26,000 citations
 - Delivered research that has had a major impact on the UK's response to COVID-19
 - c) Creating a culture of 'team science' that works across organisational boundaries
 - Brought together more than 1500 researchers across 32 organisations
 - Involved over 43,000 patients and public in leading, shaping and delivering research
 - d) Leading in partnership with Office for National Statistics (ONS) the UK Government COVID-19 Data and Connectivity National Core Study (co-sponsored by Government Office for Science and UK Research and Innovation (UKRI))
 - Made 91 COVID-19 datasets available for research across five national TREs (ONS, NHS Digital, Scotland, Wales, Northern Ireland)
 - >315 active projects underway involving >640 researchers and produced >200 papers across the UK
 - Supported policy response including multiple presentations to SAGE, SPI-M, JCVI, MHRA, Scottish and Welsh Governments
2. This response draws on experience and insight from across these initiatives and from broad discussions and engagement with stakeholders across the health data research community.

The potential benefits, including to research, to effectively use and share data between and across Government, other public bodies, research institutions and commercial organisations

3. The UK's science and research impacts are amongst the best in the world and the UK has extraordinary data science capability. Universities feature prominently in global rankings for teaching and research. Institutes such as the Alan Turing Institute, Wellcome Sanger Institute and European Bioinformatics Institute, combined with lifelong NHS health records for 67 million people and initiatives such as UK Biobank, Genomics England, National Institute for Health Research (NIHR) Biomedical Research Centres, NHSX and NHS Digital, ADR UK, Our Future Health, CPRD, OpenSafely and the capabilities of the Office for National Statistics (ONS), collectively offer strong foundations for delivering research insight and impact.
4. During the pandemic the world has engaged with data and information more intensively than ever. Data research has played, and continues to play, a pivotal role in supporting the response to the COVID-19 pandemic. For example, through:

- Allowing the NHS to identify and protect millions of people at high risk of complications from COVID-19
 - Delivering and monitoring the COVID-19 vaccination programme and allowing vital research into the safety and benefits of vaccines
 - Recruiting people to the world’s largest clinical trial of treatments for COVID-19, the RECOVERY trial, which has given the UK and the world reliable evidence about which treatments do and do not save lives and improve outcomes among patients in hospital with COVID-19. Just four months into the pandemic it was discovered that dexamethasone significantly reduces the risk of death for people hospitalised with COVID-19. This was enabled by safely and securely analysing data from hospital records that were provided by NHS Digital.
 - Enabling researchers, the public and policy makers to estimate excess deaths from COVID-19 through an online risk calculator
 - Providing an understanding of the indirect impacts of COVID-19, including the effect on cancer services and on people’s mental health
 - Enabling an understanding of the disproportionate impact of COVID-19 on ethnic groups
 - Linking healthcare data across England, for more than 54 million people, including COVID-19 test data, GP records, information about hospital admissions, with data on heart disease and strokes, enabling researchers to understand the likely impact on people with heart disease
 - Tracking how COVID-19 spreads through discovery and secure access to anonymised data from millions of users of the COVID-19 symptom study app via the Health Data Research Innovation Gateway.
5. These are just some examples of how multiple types and sources of data can be linked to enable research insights that are improving people’s lives. There is an opportunity to build on the lessons we have learnt during the Covid-19 pandemic to address any new and future health challenges. This, however, will require streamlining of processes to enable fast data access and linkage for urgent and critical research.

The barriers to sharing of health data for research

6. The transformative potential of health data research in the UK is far from being realised in full. For example, there remains a significant gap in the inclusion of primary health care data in research studies in England. Conditions that are managed mainly or exclusively outside of hospital settings are systematically under-represented in large-scale observational and interventional research studies. The result is that people with conditions such as arthritis and other causes of pain, dementia and other neurodegenerative conditions, impaired vision or hearing, many respiratory conditions, heart failure and mental health problems who are generally cared for by their GP, are missing out on opportunities to benefit from vital research.
7. There also remains a gap in understanding the longer-term complications of COVID-19. The availability of primary care health data for research studies would enable analysis of treatments that may be effective in severe cases for those with for ongoing symptoms, such as breathlessness, palpitations and fatigue. Data must be made available to understand the impact of the burden on the health system around waiting lists for investigations and treatments such as hip replacements and endoscopies. Data must be made available, safely and securely, to answer these important questions that are damaging people’s lives.
8. In our quest to address inequalities, increase resilience against future health shocks, and tackle climate change it is critical that NHS ‘real-world’ data can be linked with other diverse and large-scale data

including geospatial, multi-omics, imaging, pathology, environmental, social, administrative and citizen-generated remote monitoring data assets. This data is held by multiple data custodians. An effective and scalable research ecosystem must enable participation of government, industry, charities, universities and research institutions.

9. There are several barriers to wider health data sharing and linkage:

Public trust

10. To enable discoveries that improve people's lives through data-driven research, patients and the public need to have trust and confidence in the safe, secure and trustworthy access to and use of their data. Ideally, patients will actively want their data to be used for research that improves and saves lives. However, the public response to the GP Data for Planning and Research (GDPR) programme and rate of national data opt-outs³ demonstrates that a significant portion of the public in England have concerns about use of their data for purposes beyond their direct care. Demonstrating trustworthiness and confidence in a complex data landscape will be challenging but is vital for the UK to achieve the potential of future research.
11. Public trust is a global challenge. The Science Academies of the Group of Seven (G7) nations reported that the value of data to respond to the global pandemic may have been more fully realised if policies were defined and implemented to engage the public in the collection, use and understanding of data, with due regard for the social and political differences between nations.⁶
12. Earning trust requires meaningful public involvement and engagement. Organisations across the UK data landscape must work together to embed public involvement and engagement and bring the public into decisions about data access and use. Evidence has shown that active and meaningful public involvement and engagement also leads to more relevant and higher quality research studies.⁷ There is also evidence that the public are assured by the five safes framework (Safe data: data is treated to protect any confidentiality concerns; Safe projects: research projects are approved by data owners for the public good; Safe people: researchers are trained and authorised to use data safely; Safe settings: a secure environment that prevents unauthorised use; Safe outputs: screened and approved outputs).⁸
13. Transparent processes for taking decisions about access to data and granting access are also vital to build public trust and confidence in the use of data. Transparency, which is one of the key principles of the UK General Data Protection Regulation (UK GDPR) and the 8th Caldicott principle⁹, must underpin all decisions about data access. Use of health data by for-profit organisations is an area of particular sensitivity and public involvement, engagement and transparency are particularly important in this context.
14. HDR UK has embedded public involvement and engagement in its work to ensure that data research is shaped and delivered based on public views. Our Public Advisory Board has made recommendations to increase the transparency of the use of health data, through publicly available data use registers, which are being championed by members of the UK Health Data Research Alliance¹⁰ and will be embedded within the UK Health Data Innovation Gateway, an open website where researchers can search, discover and request access to hundreds of datasets, tools and resources for research, and the public can see what research is being undertaken by who and for what purposes.¹¹
15. HDR UK is working in partnership with ONS, ADR UK, NHS Digital, Cancer Research UK, Research Data Scotland and our public partners to collaborate on best practice of bringing the views of patients and the public to policymakers and data custodians, to ensure that data is managed and used in ways that demonstrate trustworthiness.

Fragmented landscape and slow data access

16. There are hundreds of public sector data custodians across a complex institutional and sectoral landscape. Each data custodian is responsible for creating a process for researchers to apply for access to data they hold. The data access application, review and approval processes vary considerably between data custodians. These processes can be slow and challenging to navigate for health data researchers, with data access requests **frequently taking six months or even years to be approved**. To answer a research question typically requires access to and linkage of multiple data sets and the process is particularly complex and time-consuming where there are multiple data controllers with separate data sharing processes, committees and governance structures. In addition, even across the UK there are different processes for suspending duty of confidence to facilitate data linkage for public benefit. e.g., HSC-PBPP¹² and CAG¹³, which adds to the complexity. These issues and the challenges they present for health data research in the UK have been acknowledged by several groups^{14 15} and in multiple publications.^{16 17}
17. Simplifying and streamlining data access governance processes across data custodians from the four nations therefore has the potential to dramatically speed up data access and enable data linkage in a way that is transformational for the health data research community and patients. To realise this opportunity, **HDR UK is convening a pan-UK Data Governance Steering Group to design and promote innovative and streamlined information governance approaches that facilitate faster researcher access to data in TREs**.¹⁸ The National TREs including SAIL, ONS, NHS Digital (which will be merged into NHS England), Research Data Scotland/eDRIS and Northern Ireland have agreed to support and participate in the pan-UK Data Governance Steering group. The Steering Group will also include public and patient representatives.
18. **A streamlined UK-wide process that provides robust, transparent, consistent, appropriate and proportionate governance and scrutiny of requests to ensure the five safes principles are maintained, across multiple data controllers, would significantly improve research productivity. A similar approach was adopted by the Integrated Research Approvals System (IRAS) providing a single system for applying for the ethical permission and approvals for health and social care / community care research in the UK.**¹⁹

Data availability, quality and technical barriers

19. Only a fraction of NHS, biomedical and health-relevant data is accessible to inform research and data is of variable quality. To facilitate effective use of data, it is important to provide as much information as possible about the data itself, including information about the data structure, governance processes and restrictions on use, in advance of the access request process. This information (“dataset metadata”) reduces the frequency of inappropriate data access requests. As part of the work of the Health Data Research Alliance, HDR UK has developed an open and widely adopted specification for metadata to make this information available. The specification²⁰ is available on GitHub, and the metadata for over 700 datasets is publicly available on the Innovation Gateway²¹.
20. Improving data quality also requires data to be inclusive and representative of the diverse population across the UK. The National Statistician has set out a series of initiatives to ensure the inclusivity of UK data.²² The STANDING Together Project²³ (Developing STANdards for data Diversity, INclusivity and Generalisability) is developing standards on how AI datasets should be composed (‘who’ is represented in the data) and transparency around the data composition (‘how’ they are represented).

21. Many datasets are still held in disparate environments and/or on data platforms that lack the computing infrastructure required for advanced analysis. An increasing number of health sector organisations have started to use TREs (we comment further on TREs in paragraph 38 below). However, as this is developing area, currently TREs are disconnected, both internally across different departments of a health sector organisation and externally across the health data research landscape.

Skills shortages

22. Across all sectors, data skills are highly sought after with potentially 178,000 to 234,000 vacancies and almost half of businesses struggling to recruit.²⁴ In 2019 the Tolpol Review reported that “Within 20 years, 90% of all jobs in the NHS will require some element of digital skills. Staff will need to be able to navigate a data-rich healthcare environment. All staff will need digital and genomics literacy.”²⁵ These shortages reflect the need for more people to be trained in health data science, notably health care professionals need digital skills to support research in the NHS, and people with strong quantitative aptitude need to be brought into health. Furthermore, the skills landscape is changing rapidly, and existing health data scientists need to develop new skills. These skills go far beyond traditional statistical analysis of health datasets, and require people to have strong computational (programming, modelling, data wrangling, software engineering, data standardisation) or mathematical skills as well as knowledge of health and disease. These training needs are echoed in the Science Industry Partnership: Life Sciences 2030 Skills Strategy.²⁶
23. In response to the skills shortage, HDR UK has created opportunities which attract and build talent at all stages of the health data science career pathway. We have established HDR UK’s ‘Futures’²⁷, an online learning platform which allows users to access fundamental data skills training for free and on demand. HDR UK has also supported building deep knowledge and a specialist skills base through apprenticeships, internships²⁸, masters’ programmes, PhD studentships²⁹ and post-doctoral fellowships.

Legal concerns

24. We encounter significant difficulties with public bodies being reticent to re-share clinical data for research (including pseudonymised or “de-identified” data). There is a cultural reluctance of data custodians and researchers to share data beyond their research team. In our experience much of this reluctance comes from concerns that institutions will unintentionally breach data protection legislation and face large fines from the regulator. **Clear, specific guidance from the Information Commissioners Office (ICO) stressing that the re-use of health data sets for research purposes is lawful, subject to relevant considerations and safeguards, would assist in overcoming these fears.**
25. One additional proposal to tackle this reticence and rapidly expedite access to data for research into the global pandemic is for the government to underwrite the liability of public bodies in respect of the sharing of data specifically for pandemic health research, subject to appropriate controls and safeguards.
26. Data minimisation is a key principle of GDPR and is necessary to mitigate against re-identification of individuals and to demonstrate responsible, trustworthy, and proportionate use of data to the public and patients. However, implementation of data minimisation principles can present practical challenges for health data research. To demonstrate compliance with data minimisation principles, data custodians typically require that researchers identify the specific set(s) of data required to answer their research question, which can inhibit hypothesis-free research. Where data is held in a TRE (see

paragraph 38), data custodians may be required to create a unique, curated subset of data for each individual research project (rather than allowing access to all data held in the TRE), which can lead to delays in data access. Further consultation is needed, with meaningful public and patient involvement, to consider how data minimisation principles should apply to data held in a TRE or other secure research environment. The application of these principles should ensure that vital research is not delayed or inhibited, whilst maintaining public trust in responsible and proportionate use of data.

Challenges with international data transfers

27. Research is global. To realise the benefits to the public from health data research, international data transfers and linkage of datasets from different countries are essential. For example, genomic sequencing data for COVID-19 from Wuhan, China, formed the basis of for the first batch vaccines developed in the UK. However, navigating the complexities of international transfers is a difficult burden for organisations to overcome.
28. The E.U. has only granted “adequacy” status to 13 countries outside the European Economic Area (EEA), some with partial findings only allowing certain data flows. Transferring to countries that are not on the adequacy list requires the data controller to complete a Transfer Impact Assessment (TIA),³⁰ a burdensome and complex process involving individual evaluation of international legal regimes.³¹
29. In addition, the “Standard Contractual Clauses” (SCCs), which are typically required to be executed before data can be transferred to a country that is not on the adequacy list, are often not fit for purpose in a health research context or not acceptable to public bodies outside the UK and EEA. The requirement to put SCCs in place can cause delays to or prevent research projects.³² It is unclear whether the derogation for “task in the public interest” under GDPR applies in an international health data research context and it would be beneficial to have guidance from the ICO on this issue.

The extent to which data issues are appropriately addressed by the Government’s National Data Strategy, its draft strategy, data saves lives: reshaping health and social care with data, and its consultation Data: a new direction

30. HDR UK strongly supports the Government’s vision to encourage responsible, secure use of health data for research. We would note the following key points which we believe are vital to enable successful delivery of the Government’s vision and strategy.
31. Public and patient involvement and engagement: The government’s National Data Strategy and the *Data saves lives* strategy paper recognise the importance of maintaining public trust and transparency. However, we would welcome further detail on how the public and patients will be meaningfully involved. This should not only be through engagement campaigns, but by bringing the public on board as partners in decision-making and shaping activities. Public involvement and engagement is a key part of HDR UK’s work as explained in paragraphs 14 and 15 above and the government has an opportunity to build on and learn from HDR UK’s leadership and experience. **We would welcome the opportunity for further discussion and input around involving public and patients in developing the government’s data strategy.**

32. Four nations approach: Whilst the focus of the *Data Saves Lives* strategy paper is England, we note and support the strategy's commitment to work across central government and across the devolved administrations to improve appropriate data linkage to support people's health and wellbeing. Co-ordination with devolved nations is vital for achieving a streamlined health data ecosystem. Improved connectivity across national and regional NHS partners in England, Northern Ireland, Scotland, and Wales is needed for health data research at scale.
33. Safe and secure data access: The *Data Saves Lives* strategy paper acknowledges the increasing role of TREs. HDR UK is committed to an approach to health data access based primarily around TREs as they support the highest standards of information governance, transparency, and security by removing the need for data to be physically shared between different users³³. We strongly recommend that efforts around the development and use of TREs build upon existing and evolving best practice and investments, for example the Office for National Statistics (ONS) Secure Research Service (SRS)³⁴ and the SAIL Databank,³⁵ rather than creating entirely new technical approaches. HDR UK is working with ADR UK to lead Phase 1 of DARE UK (Data and Analytics Research Environments UK)³⁶—a programme funded by UKRI to design and deliver a more coordinated national data research infrastructure for the UK. This is to support cross-domain, cross-jurisdictional linkage and analysis of sensitive data at scale for public good. The goal is a secure, interoperable data research infrastructure, based on collaboratively agreed and adopted verifiable standards, and aligned with wider domestic and international efforts in this space. This first Phase - 'Design and Dialogue', which runs from July 2021 to August 2022 - involves close collaboration with a range of stakeholders to gather detailed requirements from across different disciplines and sectors. Key to the programme is the development of a trustworthy data research infrastructure that maintains the confidence of the public. A UK-wide public dialogue is being undertaken to explore public interests and concerns regarding how sensitive data should be managed and made accessible for research across the UK. A final report on the findings of the public dialogue is due to be published in Spring 2022.
34. Legislative change proposals: HDR UK welcomes legal reforms that facilitate the greatest possible scientific and societal benefit from health data research, whilst protecting individuals' privacy. However, we do not believe that there is a strong case for many of the legislative changes proposed in its consultation *Data: a new direction*, as regulator guidance would address many of the areas of concern or ambiguity. We are concerned that some of the proposed changes would put the UK's data adequacy status with the European Union, which is vital for health data research, at risk. We are also concerned that the proposals could damage public trust in the use of their data. We have responded in full to this document³⁷.

Ethics underpinning the use and sharing of individuals' data in health and care contexts

35. Health data research must always be completed by bona fide researchers and if necessary subject to Research Ethics Committee approval.³⁸
36. Researchers must also be able to clearly demonstrate the public benefit. The National Data Guardian recently published guidance³⁹ on evaluating public benefit in health and social care data. It is important that processes for evaluating the benefit for the public are fully transparent.

The extent to which appropriate safeguards and privacy are applied in the usage and sharing of individuals' data

37. The five safes framework⁴⁰ is widely respected and adopted in the health data research community. In addition, there are also exemplar data principles: FAIR Framework⁴¹, CARE Framework⁴², WHO Data Principles⁴³ and Roche's Policy on Sharing Clinical Study Information⁴⁴. These principles help set out the rules, standards, tools, processes, roles and responsibilities to help access, use and share data at all levels – organisational, project, regional, national and international.
38. It is often not feasible to fully anonymise data for health research, especially where answering a research question requires linkage of multiple data sets. Privacy enhancing technologies therefore play an important role in enabling responsible research and maintaining public trust. The health data research community is increasingly using TREs to ensure the security of data. TREs provide a secure environment where researchers can access and analyse data in the TRE environment only and are unable to download the data, with all research outputs being subject to an approval process and statistical disclosure controls to ensure that they do not contain personal data. TREs therefore aim to maximise data security and protect individuals' privacy whilst enabling researchers to access rich, linked data assets, and are an important tool to demonstrate trustworthy use of data. TREs also enable remote research with researchers working across institutional and geographical boundaries. HDR UK is working with colleagues in NHSX to support their policy development around TREs.⁴⁵ There is currently no formal accreditation scheme for TREs, HDR UK believes that introducing such a scheme would promote their responsible use and the growth and the creation of a secure ecosystem of TREs. TREs must also be open to independent public scrutiny, with opportunities for public and patients to be directly involved in their oversight, including public participation in independent data access committees.

The effectiveness of existing governance arrangements, e.g., the Centre for Data Ethics and Innovation

39. The existing governance arrangements for data access are complex, time-consuming and do not enable rapid data access and linkage in many cases (see paragraphs 16 and 17). We believe that these challenges are best solved by building on existing best practice around data governance, for example the Office for National Statistics ONS Secure Research Service, SAIL Databank, and Scottish Safe Haven, rather than creating entirely new approaches. The pan-UK Data Governance Steering Group convened by HDR UK (see paragraph 17) will identify examples of best practice to build on in order to design and promote streamlined data governance approaches across data custodians and TREs.
40. The "Research Accreditation Scheme" under the Digital Economy Act 2017 (DEA) ensures that only bona fide researchers with appropriate skills and training access data. However, Public Health Authorities are unable to make data available under the DEA. The pan-UK Data Governance Steering Group will explore implementing a similar accreditation process for health data researchers.
41. HDR UK has not had any direct engagement with the Centre for Data Ethics and Innovation (CDEI) which may reflect the divide between health data and other forms of data processing such as those covered under the Digital Economy Act or uses of AI, we would welcome a streamlined approach across the entire data ecosystem to fully realise the potential of health data.

For more information please contact: enquiries@hdruk.ac.uk

References

- ¹ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>
- ² UK Health Data Alliance (2021) *New principles published to improve public confidence in access and use of data for health research through Trusted Research Environments* <https://www.hdruk.ac.uk/news/new-principles-published-to-improve-public-confidence-in-access-and-use-of-data-for-health-research-through-trusted-research-environments/>
- ³ [Medical Research Council \(MRC\)](#), [Engineering and Physical Sciences Research Council](#), [Economic and Social Research Council](#), [Department of Health and Social Care \(National Institute of Health Research\)](#), [Health and Care Research Wales](#), [Chief Scientists Office \(Scotland\)](#), [Public Health Agency \(Northern Ireland\)](#), [Wellcome](#) and [British Heart Foundation](#).
- ⁴ <https://ukhealthdata.org/>
- ⁵ <https://www.healthdatagateway.org/>
- ⁶ Royal Society (2021) *Data for international health emergencies: governance, operations and skills* <https://royalsociety.org/-/media/about-us/international/g-science-statements/G7-data-for-international-health-emergencies-31-03-2021.pdf>
- ⁷ Hoddinott, Pat et al. (2018) *How to incorporate patient and public perspectives into the design and conduct of research* <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6192439/>
- ⁸ Desai T, Ritchie F. and Welpton R (2016) Five Safes: designing data access for research <https://www2.uwe.ac.uk/faculties/BBS/Documents/1601.pdf>
- ⁹ <https://www.gov.uk/government/publications/the-caldicott-principles>
- ¹⁰ An independent alliance of leading healthcare and research organisations united to establish best practice for the ethical use of UK health data for research at scale. <https://ukhealthdata.org>
- ¹¹ <https://www.healthdatagateway.org/>
- ¹² <https://www.informationgovernance.scot.nhs.uk/pbphpsc/>
- ¹³ <https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/confidentiality-advisory-group/>
- ¹⁴ Medicines Discovery Catapult and the Association of the British Pharmaceutical Industry (2019) *Use of health data by the life sciences industry* <https://md.catapult.org.uk/wp-content/uploads/2019/10/health-data-report.pdf>
- ¹⁵ National Cancer Research Institute (2020) The researchers' experience when attempting to access health data for research <https://www.ncri.org.uk/accessing-health-data-for-research/>
- ¹⁶ Macnair, A., Love, S.B., Murray, M.L. et al, (2021) *Accessing routinely collected health data to improve clinical trials: recent experience of access*. <https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-021-05295-5>
- ¹⁷ Taylor JA, Crowe S, Espuny Pujol F, et al (2021) *The road to hell is paved with good intentions: the experience of applying for national data for linkage and suggestions for improvement* <https://bmjopen.bmj.com/content/11/8/e047575>
- ¹⁸ TREs are secure computing environments that provide remote access to health data to approved researchers for use in research. For more information see: <https://www.hdruk.ac.uk/access-to-health-data/trusted-research-environments>
- ¹⁹ Integrated Research Application System (myresearchproject.org.uk)
- ²⁰ <https://github.com/HDRUK/schemata>
- ²¹ <https://web.www.healthdatagateway.org/search?search=&tab=Datasets>
- ²² ONS (2021) *More inclusive data to inform better responses to the UK's most pressing social and economic issues* <https://www.ons.gov.uk/news/news/moreinclusivedatatoinformbetterresponsestotheuksmostpressingsocialandeconomicissues>
- ²³ <https://www.datadiversity.org/home>
- ²⁴ Department for Digital, Culture, Media and Sports (2021) *Quantifying the UK Data Skill Gap - Summary version* <https://www.gov.uk/government/publications/quantifying-the-uk-data-skills-gap/quantifying-the-uk-data-skills-gap>
- ²⁵ NHS (2019) The Topol Review Preparing the healthcare workforce to deliver the digital future <https://topol.hee.nhs.uk/>
- ²⁶ Science Industry Partnership (2020) *Life Sciences 2030 Skills Strategy* <https://www.scienceindustrypartnership.com/media/2070/sip-life-sciences-2030-skills-strategy-digital-version.pdf>
- ²⁷ <https://www.hdruk.ac.uk/careers-in-health-data-science/futures/>

-
- ²⁸ <https://www.hdruk.ac.uk/careers-in-health-data-science/work-experience/health-data-science-black-internship-programme/>
- ²⁹ <https://www.hdruk.ac.uk/careers-in-health-data-science/further-education/phd-programme/>
- ³⁰ https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en
- ³¹ https://edpb.europa.eu/sites/default/files/consultation/edpb_recommendations_202001_supplementarymeasuresransferstools_en.pdf
- ³² Several articles detail the difficulties US bodies have in signing up to the SCCs and how research can be impacted because of US government agencies unwilling to sign them.
Bradford L, Abboy M, and Liddell K (2020) *International transfers of health data between the EU and USA: a sector-specific approach for the USA to ensure an 'adequate' level of protection*
<https://academic.oup.com/jlb/article/7/1/lsaa055/5871850?login=true>
Rabesandratana T (2019) *European data law is impeding studies on diabetes and Alzheimer's, researchers warn*
<https://www.sciencemag.org/news/2019/11/european-data-law-impeding-studies-diabetes-and-alzheimer-s-researchers-warn>
- ³³ HDRUK and NHSX (2021) Building Trusted Research Environments
https://zenodo.org/record/5767586#.Ye7Mj_7P02w
- ³⁴ <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>
- ³⁵ <https://saildatabank.com/>
- ³⁶ <https://dareuk.org.uk/>
- ³⁷ <https://www.hdruk.ac.uk/wp-content/uploads/2021/11/19112021-HDRUK-response-Data-a-New-Direction-.pdf>
- ³⁸ Research Ethics Committees (RECs) review proposed research applications and review whether such proposal is ethical. For more information see: <https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/research-ethics-committee-review/>
- ³⁹ <https://www.go-fair.org/fair-principles/>
- ⁴⁰ <http://www.fivesafes.org/>
- ⁴¹ <https://www.go-fair.org/fair-principles/>
- ⁴² <https://careframework.org/>
- ⁴³ <https://www.who.int/data/principles>
- ⁴⁴ [https://www.roche.com/dam/jcr:1c46aa73-cea0-4b9b-8eaa-e9a788ed021b/roche_global_policy_on_sharing_of_clinical_study_informationV2.1%20April2020%20\(1\).pdf](https://www.roche.com/dam/jcr:1c46aa73-cea0-4b9b-8eaa-e9a788ed021b/roche_global_policy_on_sharing_of_clinical_study_informationV2.1%20April2020%20(1).pdf)
- ⁴⁵ <https://www.nhsx.nhs.uk/blogs/joining-up-the-dots-driving-innovation-research-and-planning-through-trusted-research-environments>