UK Health Data Research Alliance

Building Trusted Research Environments

Principles and Best Practices; towards TRE ecosystems

UK Health Data Research Alliance

Building Trusted Research Environments v1.0 dated 8th December 2021

UK Health Data Research Alliance

Foreword

The opportunities for data driven innovation in health and care have never been greater. The increasing availability of high-quality linked data has facilitated the identification of new diagnostics and treatments, improvement to health services and development of digital health technologies. It has also been essential in the delivery of population health intelligence, which has been clearly demonstrated in the response to the COVID-19 pandemic.

The Department of Health and Social Care's Data Strategy which sets the direction for the system's use of data, outlines a more effective use of research and analysis of NHS data, to support health and wellbeing, (including public health and social care), innovative practice and service planning. The use of Secure Data Environments (such as Trusted Research Environments) for access to data is key to this, as they support the highest standards of information governance, transparency and security by removing the need for data to be physically shared between different users. Data remains within a secure environment, is analysed in situ and only by those whose credentials have been established by an accredited authority. They also bring real benefits for those with a legitimate need to access, link and analyse NHS data, by providing tools and compute resources for scaled analysis, as well as facilitating collaboration and federation. Strict controls on what code and tools can be brought into the environment, and what data or outputs may be extracted further enhance the security of these systems.

To ensure the effectiveness of data driven insights and technologies it is essential to be transparent on these developments and demonstrate trustworthiness and build confidence in the new systems. The more we use patient data, in an ever-widening range of applications, the more people naturally seek reassurances that their information is secure and cannot be exploited. Establishing Secure Data Environments as the default route through which NHS organisations provide access to their anonymised data for research and analysis can give the public this reassurance.

Setting up these Secure Data Environments requires partnership working by NHS organisations as data controllers, tech providers, policy experts from the voluntary and public sector, analysts and researchers, and the public. NHSX is leading this process in England, convening a range of experts, including patient and public voice advisors to harness insights and experience.

HDR UK's thought-leadership and work to build an evidence base for Trusted Research Environments (TREs) is exemplified by this paper, which illustrates how safety is at the heart of the Secure Data Environment model and is critical to demonstrating trustworthiness to the public and professionals. This paper offers excellent insights, illustrating how TREs can operate effectively in the NHS, and is well-timed to feed into the outputs NHSX will deliver in the coming months as we work with HDR UK and other stakeholders, and set the policy on Secure Data Environments. In the New Year, the final version of the Data Strategy will be published, setting out the key principles for their use, to be followed by a technical specification and an accreditation framework.

UK Health Data Research Alliance

Secure Data Environments are the future, and to deliver on our commitments and to demonstrate trustworthiness to the public, it will be critical to work in partnership, and we thank HDR UK and colleagues across the UK Health Data Research Alliance for their extensive work and contributions.

Simon Madden (Director of Data Policy, NHSX), Catherine Pollard (Director of Tech Policy, NHSX)

Executive Summary

The UK Health Data Research Alliance (the 'Alliance') is an independent alliance of data providers, custodians and curators dedicated to improving human health by maximising the research potential of multiple forms of data at scale. The Alliance is committed to an approach to data access based primarily around Trusted (Trustworthy) Research Environments (TREs), a type of Secure Data Environment; with appropriate robust and independent accreditation, monitoring and auditing.

Adopting this approach is a way of addressing public concerns and enhancing public confidence in the use of health data for research in the UK, and further it aligns with the ambition set out in the Department of Health and Social Care's Data Strategy for the use of Secure Data Environments when accessing NHS data for analysis and research. For researchers, it involves a significantly different way of working with these wider health datasets (though does not necessarily change the modes of access to existing research cohort data). Rather than extracts of individual level data being 'released', TREs provide access to a secure analytics environment (i.e. a safe setting) where researchers bring analysis algorithms to the data.

To achieve this shift, it is necessary to mitigate concerns and demonstrate trustworthiness to:

- The public and patients through improved explanations of the benefits and risks associated with using health data as well greater transparency and lay descriptions of the technical solutions being implemented.
- Data custodians who must be willing to make data available for linkage and use in TREs.
- Researchers, many of whom are used to a data release model, who may be concerned at the detrimental impact on researcher productivity from working in a TRE environment.

There are now multiple examples of TREs operating successfully in this way, both for healthcare data and other potentially sensitive data, and the responses to the COVID-19 pandemic has accelerated this shift[1]. Examples include, but are not limited to:

- Scotland Data Safe Haven programme[2]
- UK Secure eResearch Platform in Wales[3]

[1] https://www.hdruk.ac.uk/wp-content/uploads/2020/04/200416-COVID19-Research-Data-Final.pdf

[2] https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens

[3] https://saildatabank.com/wp-content/uploads/UKSeRP_Brochure_v1.5.pdf

- Genomics England Research Environment[4]
- Office of National Statistics Secure Research Service[5]
- UK Data Service Secure Lab[6]
- NHS Digital TRE for England[7]
- OpenSAFELY[8]

There are a growing number of practical and cost saving benefits to this approach. It can maximise the utilisation of High-Performance Computing, whilst avoiding the costs of transferring and storing duplicates of increasingly large datasets, particularly imaging and genomic modalities. It also avoids the liabilities for researchers from having to ensure security of downloaded datasets.

The consultations on the previous green paper[9] indicated broad support for the direction of travel, especially from patient and public representatives. This document is an expression of commitment to promote the public interest in scientifically sound, ethically robust research while appropriately protecting the privacy and other interests of the people whose data are used in such research. It is an initial step in codifying principles and best practice to help develop the TRE approach, also expanding it with principles towards supporting an ecosystem of TREs.

This document is designed to serve as a good governance template. It is intended as a guide for colleagues across the UK and for others involved in data sharing and information governance both within and beyond the health sectors. It is not intended to cover exhaustively all aspects of governance, nor is it a statement of legal rules. It is assumed that all parties involved in data sharing are aware of their legal responsibilities and comply with them. Rather it follows that of the OECD Guidelines[10] in that it identifies areas of governance which are not found in law, or which require further expression and explanation as instances of good governance. As such, it contains a statement of the principles that should guide data sharing and linkage practice within TREs as well as instances of best practices drawn from the experiences of colleagues working in the Alliance and internationally.

The next steps in developing and codifying best practice for TREs and TRE ecosystems will be taken through the workstreams being jointly organised by NHSX and HDRUK for NHS data and the DARE initiative of UKRI, ADRUK and HDRUK for wider research data. This joint working will provide evidence towards the development of Secure Data Environment policy, which will encompass TREs. These policies for NHS in England will be defined and set by NHSX, part of the Department of Health and Social Care.

[4] https://www.genomicsengland.co.uk/about-genomics-england/research-environment/

[5] https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme

[6] https://www.ukdataservice.ac.uk/use-data/secure-lab/how-it-works.aspx

[7] https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england/

[8] https://opensafely.org/

[9] https://ukhealthdata.org/wp-content/uploads/2020/07/200723-Alliance-Board_Paper-E_TRE-Green-Paper.pdf

[10] https://www.oecd.org/corporate/mne/

UK Health Data Research Alliance

Alliance vision

The Alliance vision is that every health and care interaction and research endeavour will be enhanced by access to large scale data and advanced analytics.

1.Why does this matter?

Even before the COVID-19 pandemic, challenges to human health and health system sustainability have been increasing across the world. Whilst heart disease, stroke and cancer still account for nearly two thirds of all deaths globally, increasingly people are living with multiple diseases and long-term conditions. These affect us deeply – they change the lives of those with the diseases and those who care for them. By making health data available to researchers, we can develop a better understanding of these diseases and find ways to prevent, diagnose, treat and manage or cure them. It is also essential for tackling the direct and indirect impact of pandemics and other public health challenges.

2.Why is the UK the best place to do this?

The UK has some of the richest health data anywhere in the world. With the NHS it is feasible to collect longitudinal health data on a large and diverse population, and to make national-scale improvements to health and care. This has been demonstrated over the last 12 months through the National Core Studies and related developments including ISARIC-4C, COG-UK, OPEN SAFELY and the BHF Data Science Centre. Combined with unique research expertise, outstanding talent in the NHS and universities, and vibrant life sciences and technology industries, the UK has an unprecedented opportunity to use data at scale to drive innovation, grow the UK industry base and improve the long-term health of the public. For example, the RECOVERY trial[11] is currently the world's largest trial of COVID-19 drugs[12].

3.**How do we ensure this happens in a safe way that retains and enhances public trust?**

Research conducted by Understanding Patient Data and OneLondon, and findings from the National Data Guardian's dialogue on making public benefit assessments when using health and care data have identified that people are generally comfortable with anonymised data from medical records being used for improving health, care and services, and for research, provided there is a public benefit. [13] The more informed people feel, the more they are likely to support these uses. However, people are more likely to be uncomfortable with the idea of commercial companies accessing their health data, and there are concerns about information being passed on for marketing or insurance purposes. There are recent examples of projects which have been widely reported in the media and may have increased public concern about health data use[14].

[11] https://www.recoverytrial.net/

[12] https://www.bbc.co.uk/news/health-52478783

[13] https://understandingpatientdata.org.uk/how-do-people-feel-about-use-data

[14] For example: https://www.theguardian.com/commentisfree/2020/feb/16/our-personal-health-history-is-too-valuable-to-be-harvested-by-tech-giants, https://www.theguardian.com/technology/2020/feb/08/fears-over-sale-anonymous-nhs-patient-data

UK Health Data Research Alliance

Overview

Purpose

Health data used for research and innovation comes from a variety of sources, but most relates to peoples' interaction with the health and care system in some way – for example as an NHS patient, a participant in a clinical trial, being involved in a genomics initiative or as a blood donor. Therefore, achieving the **confidence and trust of patients and the public** in the use of this data is central to achieving our vision.

Our green paper[15] provided the case for Trusted Research Environments (TREs) which protect - by design - the privacy of individuals whose health data they hold, while facilitating large scale data analysis that increases understanding of disease and improvements in health and care (see Appendix A for the case for implementing TREs through safe settings). Extensive consultations have indicated broad support for this approach, especially from patient and public representatives (see Appendix B).

This document expands on the principles previously set out for implementing TREs based on the Five Safes[16] and a safe setting model. It also outlines additional considerations to enable TREs to function as part of a federated ecosystem.

The Alliance is an alliance of leading health, care and research organisations united to establish best practice to enable the ethical **use of UK health data for research and innovation** at scale[17]. The section on federation describes the planned direction of travel to operationalise a federated infrastructure for the Alliance, as an example of how federation across this 'national grid' of TREs could occur.

This document is a statement of agreed guiding principles for governance and instances of best practice for the operation of TREs. It has arisen from discussions and deliberations of members of the Alliance and developments across England, being led by NHSX and NHS Digital, and innovation and progress in Wales, Northern Ireland and Scotland. It is intended as a high-level instrument to contribute to the design and implementation of TREs, provide evidence towards NHSX's policy development on Secure Data Environments and also inform the public and stakeholders about how TREs are governed. This is a living instrument that will be developed and amended as necessary. Key sources of inspiration include the OECD Guidelines on Human Biobanks and Genetic Research Databases (which adopts the Principles and Best Practice approach), work of NHSX, NHS Digital, the National Data Guardian, and through guidance from the Information Commissioner.

[15] Trusted Research Environments (TRE) Green Paper https://doi.org/10.5281/zenodo.4594704

[16] https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/

[17] https://ukhealthdata.org/

UK Health Data Research Alliance

Recent context

Since the Green paper was published there has been substantial additional activity around and connected to the potential adoption of TREs to transform the ability to carry out research across the health system in a way that has the confidence and trust of patients, public, clinicians and researchers.

- The draft Data Strategy for Health and Social Care - Data saves lives: Reshaping health and social care with data - launched by DHSC and NHSX (June 2021) [18], which sets the direction for the future of health and care data.
- The announcement of the Department of Health review into use of health data for research and analysis[19], led by Ben Goldacre, creator of the OpenSAFELY TRE, which specifically mentions TREs (February 2021).
- Clinical research strategy 'Saving and improving lives: the future of UK clinical research delivery', which aims to embed Clinical Research across the NHS, bolstering capacity and creating a research-positive culture - through leveraging the UK's strengths in health data, scaling existing data access platforms (including a number of TREs) and creating new digital infrastructure (March 2021).[20]
- The Genome UK 2021-22 implementation plan explicitly commits Genomics England to develop a next-generation TRE to provide improved, authorised access to genomic data and other linked data, and for OLS and NHSX (among other partners) to build federated data infrastructure for genomic data (March 2021).[21]
- The Life Science Vision provides an encompassing set of commitments to make the UK a world leader in life sciences R&D. Streamlined governance approaches and robust at scale data infrastructure are required to ensure that data from multiple sources can be linked to create a consolidated 'picture' of the whole person and continuum of care pathway, identify the most suitable patients for clinical research, and continue efforts to improve quality and standardisation. The Vision details a specific intention to achieve this through accreditation of a handful of TREs - built to be interoperable and highly secure - to become the default route for accessing large-scale NHS data (July 2021).[22]
- The announcement that one of the three tests before implementing the **GP Data for Planning and Research programme** is that a Trusted Research Environment is available where approved researchers can work securely on de-identified patient data which does not leave the environment, offering further protections and privacy while enabling collaboration amongst trusted researchers to further benefit patients (July 2021).[23]

[18] https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data-draft

[19] https://www.gov.uk/government/news/new-review-into-use-of-health-data-for-research-and-analysis

[20] https://www.gov.uk/government/publications/the-future-of-uk-clinical-research-delivery-2021-to-2022-implementation-plan

[21] https://www.gov.uk/government/publications/genome-uk-2021-to-2022-implementation-plan

[22] https://www.gov.uk/government/publications/life-sciences-vision

[23] https://digital.nhs.uk/news/2021/new-plans-to-increase-protection-and-strengthen-security-for-gp-data-collection-programme

- The launch by UKRI, ADRUK and HDRUK of DARE UK (Data and Analytics Research Environments UK), which aims to design and deliver a national data research infrastructure that is joined-up, demonstrates trustworthiness and supports research at scale for public good (July 2021).[24]
- The announcement by NHSX of a programme to bring together partners across the health system to detail the role of Trusted Research Environments (TREs) in the health and care system, the standards they must meet and policies to govern their use (September 2021).[25]

With so many actors now engaged in discussion of TREs its critical to have clarity and transparency of the principles of attaining accreditation as a Trusted Research Environment to ensure trust is maintained - this paper seeks to solidify those principles.

Communication, Engagement and Public Trust

Central to gaining and maintaining trust in the adoption of this TRE approach is ongoing engagement, involvement, communication, and transparency with all stakeholders, which is emphasised in the requirements subsequently described. There has been engagement and consultation during and as part of the development of both the TRE green paper and this document which has indicated support for this TRE approach, but this is conditional on the commitment to permanent processes existing for stakeholders as follows:

Public and Patients – need to be able to access clear explanations on how the "Five Safes" protects their privacy and how this approach ensures that data remains under the control of the data custodians and not passed to private companies with the risk of use for unapproved purposes and the benefits for citizens and the UK of building an ecosystem for safe and secure health data research. This will also need to cover details of the technical controls in place, such as to protect data held in public cloud that ensures that the data cannot be accessed by the hosting organisation. TRE providers need to emulate technology companies who have enhanced trust in their systems through complete technical transparency, publishing full technical documentation of security design and implementation including assessment reports by independent reviewers and/or regulators, to allow review and discussion by both technical press and informed experts. There is valid public concern over the control of data that is made available for research and on the limits of de-identification which statements of 'trust us' will not address. Communications, engagement, and involvement of patients and public in co-design to enhance trust needs to be continuous, alongside direct involvement in the approval processes for research projects (safe projects) and ongoing transparency regarding data use and informative communication of outputs from research.

Analysts, Researchers and Innovators –need to know that their concerns about user experience are listened to and considered in the development TRE systems and that there is focus on a first-class research and innovation experience, minimising the impact of TRE restrictions on research efficiency. There will need

[24] https://dareuk.org.uk

[25] https://www.nhsx.nhs.uk/blogs/joining-up-the-dots-driving-innovation-research-and-planning-through-trusted-research-environments/

to be ongoing communications to highlight other longer-term benefits such as more rapid access to data and improved opportunities for linking data that has until now been restricted due to the data custodians risk positions.

Data Custodians – needs to be assured that TREs are secure for the data they manage and that it will remains within their control and meet GDPR and Common Law Duty of Confidence requirements. Given the enhanced level of control, compared to the data release model, there should be facilitation of discussion around how data access management processes might be adjusted to facilitate quicker access to richer data given the reduced risk of inappropriate disclosure or use.

Funders – TREs offer funders of research a range of potential benefits. These include more efficient research through improved utilisation of storage and compute resources; a proportionate approach to data access requests based on all the five safes; and audit trails on provenance of research outputs and data manipulation. This supports both transparency and replicability. These features may also benefit regulators. The TRE model should also provide a more cost-effective approach to supporting large compute and storage requirements, such as for AI training, as environments adopt either a hybrid or complete cloud model (safe compute) provided cloud providers' commercial models address some of the current issues for example around the costs of data egress.[26]

Principles and best practice for a Trusted Research Environment

The UK Health Data Research Alliance's Principles for Participation[27] include *use a proportionate approach to the governance of data access based on the five "safes"*. TREs provide a "safe setting" approach and a schematic example of a TRE environment is shown in figure 1.

Ensuring public trust is maintained requires TREs to adopt a common set of principles and implement corresponding requirements for each of the *five safes* to gain accreditation, despite being operated on different systems run by different organisations. These are outlined in this section for each of the *safes* in turn, as principles and best practices as if each TRE were operating in isolation. In the following section discussing a federated ecosystem, examples are given of how the implementation of some of these best practice elements might be through shared ecosystem components.

Best practice definitions given are mainly labelled for either **Data Custodian** or **TRE provider** to differentiate between those managing the data and operating the processes to safeguard it and those providing purely technical TRE infrastructure, which might be outsourced. It would be expected that anyone functioning solely as **TRE provider** would be technically unable to access the data.

[26] https://www.gartner.com/en/documents/3939969/the-art-of-taming-data-egress-charges-in-hybrid-and-publ

[27] https://www.hdruk.ac.uk/wp-content/uploads/2020/03/200304-Principles-for-Participationv2pdf.pdf

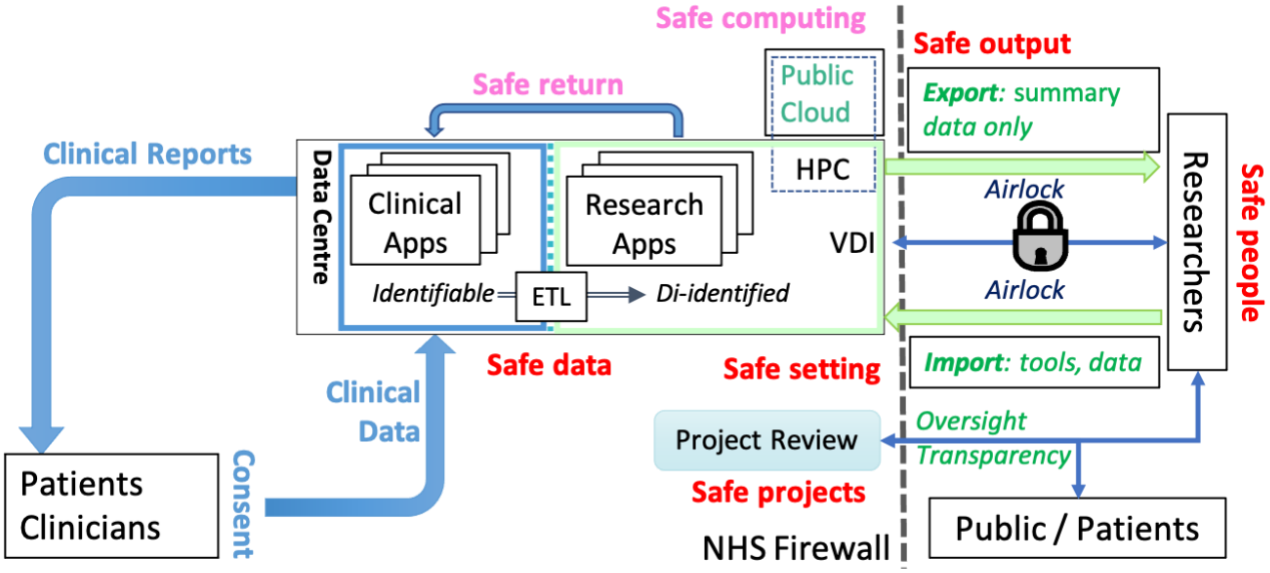

Figure 1: Schematic example of a TRE environment. The example is based on the Genomics England TRE, however labels in red identify the underpinning **five safes** that any TRE would be expected to implement. The TRE operates as a **safe setting** [green] that approved researchers (**Safe people**) can only access via a virtual desktop interface (VDI). Only de-identified data is accessible by researchers (**Safe data**). Research is overseen by a review of projects (**Safe projects**) involving ***participants***, with publicly accessible lay summaries providing ***transparency***. Only summary data can be exported from the TRE through an ***Airlock*** and only after manual review (**Safe output**).
This figure also shows two optional TRE characteristics: Analysis of large and complex datasets, such as whole genomes, requires High Performance Compute (HPC) resources, however scaling HPC to meet the needs of large numbers of researchers is challenging. This can be addressed by using public cloud resources, however implementation needs to meet the security requirements of **Safe computing**. In the Genomics England environment, the data accessed by researchers is a de-identified version of real-world data being analysed to produce reports for clinical care [blue]. This close coupling makes it possible to pass back research results that may be relevant for individual clinical care (**Safe return**) - see below.

Safe people

Individuals allowed access to TREs should be analysts and researchers[28] able to demonstrate appropriate credentials and be undertaking approved (safe) projects consistent with the requirements for access to the contained datasets. Where individuals are accessing TREs as employees it is expected their organisations (whether academic, industry or NHS) should be prepared to take responsibility for their actions and vouch for each individual. However access should also be possible for researchers from non-standard backgrounds, such as those bringing distinctive data science capabilities from other sectors such as social sciences, finance or from start-ups., provided there is transparency regarding who is accessing the data and

[28] Analysts and researchers could be from academia, NHS and from industry. Any accreditation of research must cover all these communities.

what their backgrounds and interests are to maintain trust and perception of conflicts of interest. Researchers should be required to have signed generic legally binding terms of use which could include:

- not trying to re-identify individuals[29]
- immediately reporting any security weakness found when using a TRE system and not attempting to exploit it
- not sharing their login credentials with any other individual
- informing a TRE if they are changing institutions before they have done so

They should also undertake information governance training and, potentially, training specific to the health domain, TRE and/or datasets, refreshed periodically.

Accredited TREs should have systems to track individuals and organisations, the status of their on-boarding progress.

Best Practice for SAFE People (Analysts): Only trained, authorised individuals can access the data

A1. Data Custodians should have processes to verify the identity of *Individuals* from the information they supply
A2. Data Custodians should have processes to verify the status of *Individuals* from information *they* or their *Organisations* provide regarding their Accredited / Approved / Bonafide analyst / researcher status or equivalent.
A3. Data Custodians should have processes to verify that *Individuals* undertake and renew Information Governance Training in support of their accreditation status
A4. Data Custodians should maintain records of signed agreements by *Individuals and Organisations* documenting their legal undertakings governing their access to the data, including capturing declarations of funding/sponsorship, commercial interests and any potential conflicts of interest.
A5. TRE providers should be able to apply authorisation policies to enable access to services based on the access approved for each individual
A6. TRE providers should maintain a record of all user access performed by *Individuals* for audit purposes and to generate transparency information on usage.
A7. Data Custodians should have processes to disbar users in breach of service with an appeals process

Safe projects

Despite the privacy protections offered by TREs, it remains essential to ensure that the use of data is appropriate and has the potential for public benefit, so individuals applying for access need to justify this. Applying for access is a significant undertaking, so Data Custodians should make the process as clear and transparent as possible. Data Custodians also need to provide ways for potential applicants to discover if the datasets they contain are appropriate for the proposed project (e.g. have sufficient relevant data) before making a full application, such as providing basic externally accessible summary statistics, or limited

[29] Data Protection Act 2018 Section 171 - Re-identification of de-identified personal data

query interfaces or manually responding to specific queries. Involving representatives of patients and public in the review of projects is a key element, alongside transparency of decision making and data use, as highlighted in the Foundations of Fairness joint report from Understanding Patient Data and the Ada Lovelace Institute[30] and subsequent learning data governance model proposals[31]. This includes recommendations to address the issue of frequency of feedback and transparency around the research that is being carried out, that patients, public and cohort research participants frequently raise in focus groups. To improve transparency, Data Custodians should require lay summaries to be provided as part of the project approval process with these being made public on approval.

Project proposals will frequently have dependencies on external data and software, requiring support for installation and or user importing (see requirements for Safe Setting).

Best Practice for **SAFE Projects:** Use of data must be appropriate, which projects must justify

P1. *Individuals and organisation* should provide detailed project descriptions including project methodology, funder/sponsor information, ethics approvals and time period of access
P2. Data Custodians should provide detailed guidance of the data access request process, including time frames, requirements and decision making process
P3. Data Custodians should consider providing the ability for Individuals to submit and process basic enquiries of the data they hold prior to submission of a formal access request.
P4. Data Custodians should provide a proportionate data access request form to collect all relevant information about the individual/organisation's project
P5. Data Custodians should inform and update individuals on the status & processing times of their application and allow for individual appeals process
P6. Data Custodians data access decision making processes should have meaningful involvement of patient and public / lay representatives.
P7. Data Custodians should maintain a public Data Use Register with the aspiration that it is updated in 'real time' with approved projects

Safe setting

As summarised in the appendix, there are multiple existing platforms providing research access to health data through the implementation of a safe setting. While one side of operating a safe setting is the need to ensure public and data controller trust through security and transparency, the other side is the need to ensure it is engineered to be as easy to use for research as possible.

At minimum a safe setting should implement:

[30] https://understandingpatientdata.org.uk/what-do-people-think-about-third-parties-using-nhs-data

[31] https://understandingpatientdata.org.uk/news/new-approach-decisions-about-data

- A system to hold data securely such that individual level data cannot be exported. For transparency the security design and implementations should be independently audited with reports reviewed by a patient/public oversight group and made public.
- Systems to allow secure remote access by accredited researchers to carry out analysis with the ability to keep track of researcher activity (to ensure compliance with "safe projects") and that ensures accounts cannot be shared (to ensure compliance with "safe people").
- A research environment containing a set of tools to allow data to be analysed, with a barrier between the safe setting environment and the outside world to prevent data egress.
- Processes and systems to support limited export of data as results (Safe output, see below); to support data import (Safe data, see below) and software import.

There are two main alternative approaches to provide researchers with remote access to a safe setting. The first alternative is to provide a Virtual Desktop Interface (VDI) which allows researchers to login to a system where they can view data but cannot export it or cut and paste it out of the system. Such a system makes it easier for researchers to explore raw data (always de-identified) and identify errors affecting their results, however because there are multiple interfaces including the command line it's impossible to perfectly record all analysis actions sufficient to demonstrate exactly what questions have been asked of the data. As an example, Genomics England implements a VDI.

The second alternative is to provide an interface which only allows the execution of algorithms where researchers cannot view raw data directly. This makes it harder for researchers to debug problems and generally requires provision of some dummy datasets in identical formats to allow algorithms to be tested, however it does result in a system where all actions can be exactly logged, i.e. exactly what version of code was run on what version of data at what time. This allows Data Custodians to report exactly what analysis a researcher has or has not carried out, which aids transparency and public trust. As an example, OpenSAFELY implements an interface that only allows the execution of algorithms.

This second approach also inherently supports the reproducibility of analysis since the entire analysis must be defined in code which can then be shared, and its versions tracked. The use of Reproducible Analytical Pipelines (RAP)[32] - open source of code utilised by a number of analysts through code sharing platforms - is being increasingly used to generate standardised outputs, including across UK government departments. These standard outputs are usually combined with commentary. Their use minimises the manual steps within analyses, as well as supporting audit of analyses through the use of version control software. Quality assurance can also be embedded into code that is run, using logging and automatic testing, and identifying issues in the underlying data, e.g. outliers. The use of RAP can therefore improve the efficiency and quality of analysis, while supporting the drive for transparency.

Regardless of the mechanism for providing access, there will be many cases where the types of analysis that researchers wish to carry out will go beyond statistical packages provided as standard by TREs (e.g. R-Studio, Stata) or other installed software. It is therefore important for researchers to be able to request additional packages be installed and / or have the ability to import their own algorithms.

[32] https://gss.civilservice.gov.uk/reproducible-analytical-pipelines/

The most programmatic way of importing code directly is via github, or dockerhub which requires allowing connections outside the safe setting to external services. If configured in a conventional way this would risk allowing data to be exported, however it is possible to configure intelligent firewalls to allow individuals to login to these external sites, whilst maintaining readonly connections only allowing import of code or docker containers into the safe setting.

Best Practice for **SAFE Settings:** Analysis platform actively minimizes the risk of unauthorised access, use or disclosure

S1. TRE providers should implement processes and systems that hold and managed data securely, encrypted at rest with data encryption keys exclusively controlled by Data Custodians
S2. TRE providers should implement mechanisms to provision a minimised dataset bespoke to the individuals request encrypted with a separate key accessible by the project individuals
S3. TRE providers should provide a secure environment to allow individuals to perform their analysis using tools supplied by the TRE provider and/or tools requested to be deployed by the individual
S4. TRE providers should provide services that allow individuals to remotely execute analysis workflows using TRE supplied tools or research software with minimal hands-on access to the data
S5. TRE Providers should publish their security design and implementation reports for review
S6. TRE providers should provide assurance statements that ensure their processes and systems are conformant to secure data processing standards – ISO 27001, IGToolkit/DSPT, ONS/UKSA Accredited Processor
S7. TRE providers should allow individuals to specify software, research code, reference data, configurations to be deployed with their SAFE Setting which may be subject to a review process before deployment
S8. TRE providers should make every attempt to support the ongoing collaboration between project members, including provide collaboration software – Git, Shared doc.

Safe computing – an extension of Safe setting

Since the ONS definition was originally developed, a new issue has become important that is not explicitly covered by the "Five Safes" but should be addressed to build public trust. This is the outsourcing of provision of computing infrastructure for all or part of a safe setting to third parties through partnerships with commercial organisations or use of public cloud computing providers.

Previously, safe settings have been almost exclusively provisioned through "on-premise" computer hardware where physical security of equipment, network security, software maintenance etc. is the responsibility of the data custodian or TRE provider. Such systems can be configured as "private cloud" to support the use of software distributed as virtual machines and containers. However, use of third-party computing resources, such as public cloud, offers many potential advantages for TRE providers and is likely to be the default from now onwards. This provides dynamic scalability of compute to enable short periods of intensive computation such as for AI training. Outsourcing layers of the hardware and software stack

which have become commodities to cloud providers brings other potential benefits due to their greater capacity to engineer scalable platforms and implement robust security.

In order to build public trust, use of private sector computing infrastructure to provide a safe setting should be done in such a way that none of the hardware and software layers outsourced make it possible for the third-party provider to access individual health data. This needs to be enabled through security design and engineering as well as contractual arrangements with the third-party provider to minimise the risk of a data security breach. It is accepted by some cloud providers that a security design that ensures they have no data access is a critical requirement to many organisations. Technical papers have been published about how to engineer this level of security where cloud provider administrators have no ability to access any customer data[33,34].

The security engineering and design required to make this possible is complex and involves encryption at rest of all health data and encryption key management infrastructure configured such that only the data custodian controls the keys. It is proposed that TREs using public cloud should be engineered in this way and would be regarded as operating a "safe setting" that implements "safe computing".

Explaining this complex engineering and design in ways that data custodians, researchers and members of the public can understand and engenders trust represents a challenge that needs further consideration.

Best Practice for **SAFE Computing:** Use of cloud conditional on providers being *unable* to access data.

C1. TRE providers should implement processes and systems that hold and managed data securely, encrypted at rest with data encryption keys exclusively controlled by Data Custodians
C2. TRE Providers should publish their cloud security design and implementation reports for review.

Safe data

With access to data restricted to within a safe setting, it's particularly important that the data assets held within it are as easy to understand and use as possible, to minimise the manipulations and investigations required just to understand the data, particularly if the safe setting only allows programmatic access rather than through a VDI. Comprehensive metadata descriptions of the data assets held (including data standards, vocabularies, and data profiles), including the source of the data and the lawful/ethical basis for collection and use, greatly aid discoverability and use. The impact of this can be seen from the success of the Health Data Research Innovation Gateway (the Gateway) in providing a searchable interface to all Alliance metadata (see below). Critical to this has been the development of metadata standards and

[33] AWS: https://d1.awsstatic.com/whitepapers/using_aws_context_nhs_cloud_security_guidance.pdf; https://d1.awsstatic.com/whitepapers/Security/AWS_Security_Best_Practices.pdf; https://d0.awsstatic.com/whitepapers/AWS_Securing_Data_at_Rest_with_Encryption.pdf

[34] Google: https://services.google.com/fh/files/misc/handling_healthcare_data_uk.pdf, https://cloud.google.com/solutions/setting-up-a-hipaa-aligned-project

adoption across the alliance. It is equally important that TRE Providers make their data available using appropriate data standards, (see Data Standards Green Paper[35]).

Although data is held in a safe setting, it's still critical that all direct identifiers should be removed from all data and replaced with uninformative pseudonyms to prevent accidental re-identification. Naturally it is a requirement that all data held within the safe setting should be encrypted at rest and in transit as already specified (Safe setting, **S1**). Where possible TRE providers should provide data linkage services, such as to handle cases where there are requests to import additional linked data (subject to appropriate consents and permissions being in place). Because all these processes for transforming and handling data are critical for protecting privacy TRE providers need to ensure they are comprehensively described in lay terms for the public.

Controls on access to data in the TRE should be proportionate to the approved use, with appropriate data minimisation applied, however because data assets are held in a safe setting and any export is subject to very strict data minimisation through an airlock process (see safe outputs, below) it is possible to consider allowing access to a broader set of data within the TRE than would have been historically provided as an export. The benefit is to open up new research opportunities by facilitating hypothesis-generating research within broader boundaries of a project approval, with the potential for greater public benefit. Environments such as at Genomics England, which implements a safe setting with strong implementation of the other safes, allow researchers to analyse across the entire dataset, facilitating broad investigations of genome/phenome relationships, multimorbidity effects etc. At the same time, it is also important for TREs to have the capability to create secure project specific spaces within the safe setting when required for specific projects, such as when other sensitive data is imported.

Best Practice for SAFE Data:

D1. *Data Custodians* should provide descriptive, semantic and technical metadata about their datasets publicly available in human and machine-readable form
D2. *Data Custodians* should provision data using a standardised format supporting well-known data standards, e.g. See HDR UK Data Standards Green Paper
D3. *Data custodians* should ensure all direct identifiers are removed from source data assets prior to onboarding into the TRE and provide a lay summary of how these processes are managed.
D4. *TRE provider* should provide mechanisms and process for researchers to request ingress of external (additional) data to be used by researchers as part of their research
D5. *TRE Providers* should provide data linkage services to allow users to request linkage of datasets with data held internally or externally to the TRE provider
D6. *TRE providers* should implement appropriate data minimisation proportionate to sensitivity and the approved use of the data
D7. *TRE providers* should be able to provision project specific workspaces that maintain the integrity of the provisioned data and ensure multi-tenant security and privacy

[35] https://www.hdruk.ac.uk/wp-content/uploads/2021/06/210622-Recommendations-for-Data-Standards-2021-Interim-Paper.pdf

UK Health Data Research Alliance

Safe outputs

As outlined in Safe setting, TREs must implement a barrier (or "air lock") between the safe setting and the outside world to prevent unauthorised data export (or import). TREs must implement processes and systems to allow approved data to cross this barrier. Concerning data export, this should be based on the principle of minimising the amount of data, such as that necessary to report results in a publication but no more. TREs should implement systems with functionality to track requests and decisions, supporting cycles of rejection and revision.

Data Custodians should employ airlock managers to review requests to export data with mechanisms with support from an oversight group. To minimise the risk of applications needing to be revised there should be sufficient information and training for individuals on what exports are possible and how to appropriately summarise and minimise data. TREs should ensure there is sufficient capacity to support this function else it risks being a bottleneck for users and should implement and publish KPIs on speed of turnaround etc. Equally, the integrity of this process is critical to maintain public trust and it should be expected to report on export criteria and activities to lay representatives. TRE airlock managers should aim to share expertise and develop consistent approaches to definition of safe summary outputs and consider joining the Safe Data Access Professionals (SDAP) working group[36] and consider adopting training programmes such the ONS Office for National Statistics Output checker training.

For situations where a processed dataset cannot be exported because it is too disclosative, but needs to be referenced in a publication, a possible solution is for TREs to implement systems to allow such datasets to be stably archived within the safe setting, with an identifier that can be referenced within the publication. Linked to this TREs should be able to support temporarily access to such datasets by publication reviewers if requested, as well as allowing access to researchers wanting to carry out follow on analysis within the safe setting. Similarly, where an individual has completed a project, but wants to preserve intermediate datasets that cannot be exported for a defined period, TREs should provide an archiving function to support this.

Best Practice for SAFE Outputs: Only non-disclosive output data is subject to release from a TRE

- O1. Individuals should be able to apply for a data or code release from the safe setting
- O2. Data Custodians should implement timely processes and systems to assess and decide on the data release applications in a consistent manner, including decision provenance & appeals process.
- O3. Data Custodians should provide open and clear documentation of the statistical disclosure control policies including the assessment criteria that are used to evaluate data export requests
- O4. TRE Providers should provide software solutions to manage data export requests.
- O5. TRE Airlock managers should aim to harmonise and coordinate output checking and data release management processes with other TRE Airlock managers and external expert groups.

[36] https://securedatagroup.org

O6. Data Custodians and Individuals should ensure appropriate training is afforded to staff and individuals to ensure individuals are able to produce outputs that require minimal effort to check
O7. TRE Providers should consider providing a mechanism to stably archive datasets that support publications but cannot be exported with externally visible IDs and support reviewer access if requested.
O8. TRE Providers should provide a mechanism to archive an entire project workspace for a determined duration.

Safe return - Optional extension to the TRE definition

While health data held within TREs is de-identified for research purposes to guard against accidental re-identification of individuals by researchers (safe data), there are differences as to whether it is consented and/or technically possible to send individual analysis results back to the clinical setting that originated the data and where identities are known. This would be for individual clinical care purposes or to enable invitations to be sent to specific individuals inviting them to participate in trials and other research projects. This could be codified as an optional extension to the TRE definition.

For example, in the Genomics England case (see figure 1), there is ethical approval and patient consent to pass analysis results for an individual generated in the research environment safe setting back to the clinical setting for re-identification, evaluation and return for clinical care. These 'outputs' supplement the results already generated by clinical analysis pipelines in the clinical setting. Given that the clinical analysis pipelines only produce diagnostic results in 20-25% of cases, there is considerable clinical value in additional individual diagnoses being proposed for undiagnosed patients from the research side. Making this possible requires completely robust and certified data paths for individuals to ensure that a result obtained in a research environment is always perfectly mapped back to that individual's clinical record.

Safe return would only be possible where there is appropriate ethics, consent and certified data paths between clinical and research systems. For example, this would not be possible in the case of research cohorts such as UK Biobank, which does contain clinical health data for each individual, but where there is no consent for return of results to individuals, so a TRE based on UK Biobank data could not implement this.

For TREs where return of results is possible, there can be multiple benefits. It may be only a part of the research activity carried out within a TRE, but supporting this option has the potential of increasing the convergence of research and clinical care, bringing researchers and clinicians closer together. It may also provide an additional incentive to clinicians to ensure the clinical data they record is as complete as possible if research use could result in additional clinical feedback.

Best Practice for SAFE Return: (optional, only where ethics, consent allow this)

R1. TRE providers would need to implement processes and systems that track individual consents and withdrawals held in clinical systems to ensure that Safe return is only used for those who have specifically and currently consented to it.

R2. TRE providers would need to ensure data paths from clinical to research and back to clinical are certified to ensure data mismatches cannot occur.

UK Health Data Research Alliance

Principles and examples towards a federated TRE ecosystem

To maximise the potential of using TREs, common agreed specifications and systems are needed to simplify processes for researchers, lowering barriers to access multiple TREs and supporting federated analysis.

This is an area of active work, particularly around the development of practical implementations of federation. Funders are supporting technical proof of concept development work as part of initiatives, e.g. DARE Sprint Exemplars which are currently running[37].

This section therefore presents principles that have been developed so far, using the Alliance Gateway as an example of implementation. Specific components either under development or planned are presented as part of the overall plan for a federated ecosystem of TREs across the Alliance, built around the Gateway project, again organised around the five safes definitions.

Since 2019 HDR UK has been developing the Gateway[38] as a central portal for the Alliance. Initially the only requirement for Alliance members has been to submit metadata descriptions of each of their datasets to the Gateway, which are indexed. The result is a searchable repository of more than 600 UK healthcare datasets, which has greatly increased discoverability by researchers and is being widely used. The Gateway contains no actual data and can only direct researchers to custodians of datasets of interest, many of which operate TREs. However the Gateway interface includes a login system which can federate with existing institutional academic or UK health service identity systems. This is a starting point for centrally supporting processes that simplify interactions with multiple TREs and a general model for TRE ecosystems.

In the context of the Alliance, Figure 2 shows a schematic of how it is planned that a subset of the previously listed requirements related to each *safe* for individual TREs could instead be implemented through the Alliance Gateway as a service supporting multiple Alliance TREs. This would require TREs to implement and expose a range of standardised APIs to connect to Gateway services. The Gateway team has a roadmap for implementing these services in collaboration with Alliance TREs.

[37] https://dareuk.org.uk/applications-for-our-sprint-exemplar-projects-have-closed-whats-next/

[38] https://healthdatagateway.org/

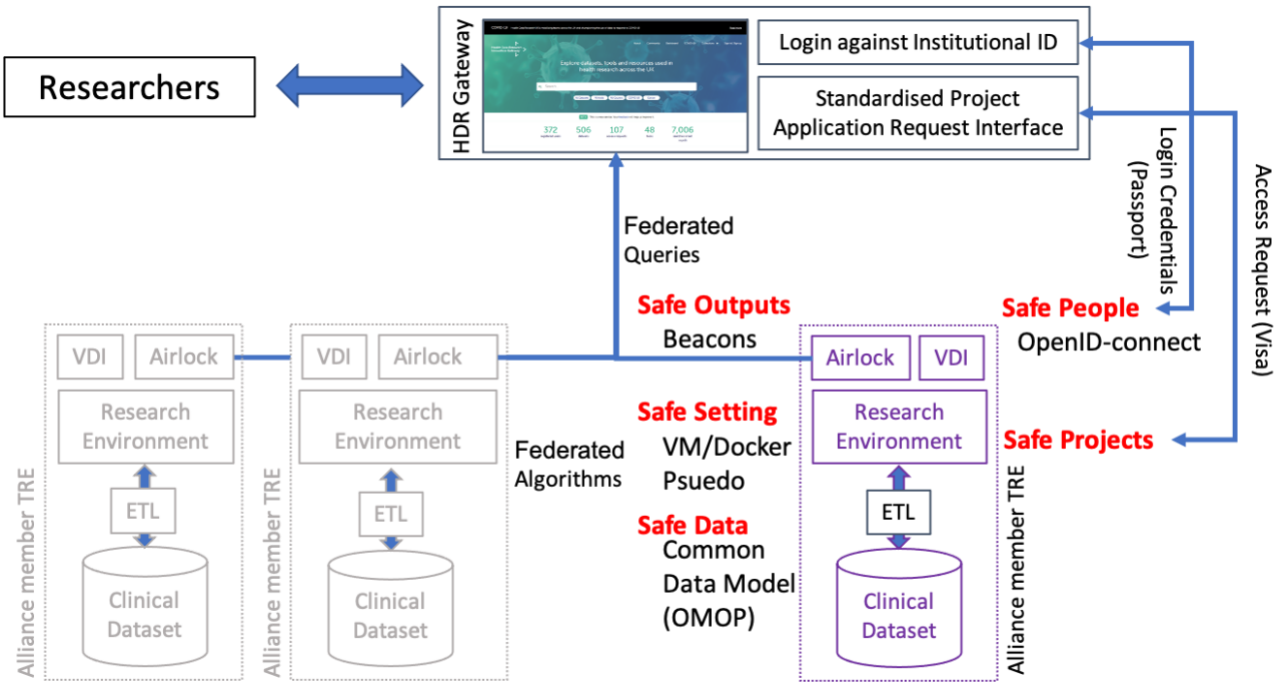

Figure 2: Schematic of additional requirements for TREs to operate in a federated ecosystem. A simplified schematic of a TRE (c.f. as shown in Figure 1) is shown in purple, with the *five safes* labelled in red. Currently the Alliance operates the Gateway with user login functionality, which provides a query interface to search metadata supplied manually by Alliance members describing each of their datasets. 1.TREs that implement an OpenID-connect adaptor can delegate requirements to implement **safe people** to the Gateway, allowing the same user credentials to be shared and the user account maintenance burden to be reduced. 2.TREs that support a standardised Access Request API can reduce their **safe project** administration and benefit from receiving requests directly from the Gateway interface. By linking project approvals to OpenID-connect credentials they can simplify user management further. 3.By implementing ETLs that transform health data elements (even partially) into common data model representations (e.g. OMOP) TREs enhance the ease of use of **safe data** within their TRE and the chance that algorithms can be reused without customisation within different TREs. 4.Support for programmatic importing of algorithms as code e.g. GitHub/Docker containers/Virtual Machines into the **safe setting** allows the same algorithm to be distributed to multiple TREs with different populations. Support for standardised one time use psuedoIDs for individuals allows algorithms to link datasets for joint federated analysis. 5.Support for **safe output** approval process and deployment of algorithms that programmatically implementing safe output rules, e.g. Beacons that can be externally queried returning replies that conform to statistical disclosure control policies - allowing programmatic export of meta-analysis summaries output from multiple TREs as well as supporting federated queries within Gateway.

There will be benefits for all stakeholders from this arrangement. It will simplify processes for individuals wanting to gain access multiple TREs by providing a common digital identity across different systems once access has been verified. For Data Custodians it can reduce the costly and frequently manual processes connected with account management and approval and make it easier to support more complex cross dataset analysis, without impacting on the independence of custodians to review project justifications and control access.

UK Health Data Research Alliance

Federation for Safe people

Of the requirements for Safe people, the verification of identity of individuals applying for data access (**A1**) is the clearest example where a centralised process should be able to serve the requirements of multiple TREs. As discussed, the Gateway has already implemented a login system which can federate with the identity systems of Academic and NHS organisations. Since those organisations will have already verified identity as part of the process of issuing logins on their systems, it should be possible for TREs to trust these identities. All that is then needed to technically implement **A1** is a link between TREs and the Gateway to federate this identity information. This is the planned approach for the Alliance ecosystem. The recommendation is to use existing industry standards such as OpenID Connect (OIDC) and OAUTH2 as a mechanism to transmit user identity information safely and securely leveraging existing institutional Single Sign On and Identity brokerage solutions such as UK Federation and NHS Identity which are used by Academic and NHS organisations receptively.

The second requirement for Safe people, to verify an individual's eligibility for data access (**A2**), is more complex to centralise because different data custodians will have different views on what experience is necessary for user accreditation. Our recommendation is to support the use of existing accreditation frameworks such as the "ONS Approved Researcher Scheme"[39] or "Accredited researcher under the Digital Economy Act 2017"[40] (provided by the UK Statistical Service Authority) to enable Individuals to undertake appropriate training and obtain an accreditation status. Given the wide range of different data sets across the Alliance and beyond (e.g. the UK Trusted and Connected Data and Analytics Research Environments, DARE UK programme, which include UKRI and ADRUK as well as HDR UK[41]) a single accreditation standard may not be sufficient to satisfy all data custodians or for access to all data types, but even if such accreditations are only accepted by a subset, this will simplify processes for many. To technically implement **A2** requires allowing digital assertions of accreditations to be linked to identity (**A1**). The Gateway plan is to extend the existing OIDC and OAUTH2 standards using the established Global Alliance for Genomics and Health (GA4GH) Passports and AAI standards[42]. More broadly, this will allow user identities to interoperate within a federated ecosystem of services and for a user to be uniquely identified by each service participating within the ecosystem. Conceptually it can be thought of as the identity (A1) acting as a 'passport' for individuals with the GA4GH extensions allowing 'visas' issued by different authorities to be attached (such as to satisfy requirement A2) and used to authenticate accreditation status and obtain authorisation to access services provided by different data custodians.

---

[39] https://researchaccreditationservice.ons.gov.uk/ons/ONS_Homepage.ofml

[40] https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#becoming-an-approved-researcher-through-the-ons-approved-researcher-scheme

[41] https://www.hdruk.ac.uk/ukri-trusted-and-connected-data-and-analytics-research-environments-programme-updates/

[42] https://www.ga4gh.org/ga4gh-passports/

With these technical systems in place, it's possible to consider digitally accepting an externally recognised Information Governance Training accreditation as sufficient to meet that requirement (**A3**), or part of that requirement if there is data set specific IG training also required.

Similarly, some of the declarations required for transparency and avoidance of conflicts of interest (**A4**) may be associated with an individual's account on the Gateway rather than needing to be captured by each Data Custodian separately, however signed legal undertakings may still need to be directly collected by individual TREs.

Finally, through this mechanism, it will be possible to communicate information about an individual across the ecosystem, such as if an individual has been disbarred from a service (**A7**). It is important that this sort of information should be shared across the ecosystem as part of maintaining public and patient trust.

Best Practice for **Federated SAFE people (Analysts)**: (Exemplar from Alliance / Gateway plans)

FA1. TRE providers should implement OIDC and OAUTH2 APIs to link to identities provided through the Gateway.
FA2. Data Custodians should define institutionally validated identities they will trust as satisfying requirement A1
FA3. Data Custodians should define which external accreditation status attached to an individual's digital identity they will accept as satisfying requirement A2.
FA4. Data Custodians should define which external evidence of information governance training attached to an individual's digital identity they will accept as satisfying requirement A3
FA5. Data Custodians should collect as much information about an individual including declarations of potential conflicts of interest from HDR gateway collected records, rather than duplicate collection, to at least partially satisfy requirement A4
FA6. Data Custodians should ensure that any disbar process is flagged against an individual's identity, so it is visible across the ecosystem. They should also actively check for such flags attached to the identity of any individual accessing their TRE as part of processes associated with A7.

Federation for Safe projects

Most data custodians provide some form of Data Access Request process and these vary from using offline Word/PDF forms to fully online access request process. The Gateway is attempting to streamline the data access request process by collating the questions asked by Alliance members to build up a consolidated Data Access Request Form as a standard Five SAFE approach to allow individuals to submit access requests. The recommendation for new Alliance Data Custodians is use this designed Five SAFE Form as the baseline to extend and modify for their use. The Alliance will be publishing the Five SAFE form as a standard to allow reuse and interoperability between the Gateway and data custodians for processing.

There currently does not exist a single streamlined data access request management API that allows individuals to submit an access request to one or more data custodians as they are implemented by each

data custodian themselves. The Alliance aims to help consolidate these APIs into a single streamlined Data Access Request API to allow interoperability between systems.

Best practice for **Federated SAFE projects**: (Exemplar from Alliance / Gateway plans)

FP1. Data custodians should consider using the HDR UK designed Five SAFE form as the baseline for collection of data for project approval as part of satisfying requirement P4.

Federation for Safe setting

Within an ecosystem of TREs each operating equivalent safe settings with the equivalent processes and overall accreditation, there are two main approaches to supporting federated analysis: Running algorithms in multiple safe settings and combining the results or moving subsets of data between safe settings to allow specific algorithms to be run in a single safe setting. Given the justifications for adopting TREs and the declarations on not moving data (see Appendix A), the former is the ideal approach, however the second may be the more practical for some cases at least until support for federated computation is more fully developed. From a public and patient trust point of view, moving data between TREs with the same security and the same Data Custodian frameworks should be acceptable, but will require careful engagement, review and support.

Supporting the execution of algorithms across multiple TREs requires support for a number of functionalities beyond those described in Safe Setting previously. Depending on the type of analysis planned, linkage of datasets may be required across TREs making it important that common pseudoidentifier generation processes are adopted across the ecosystem. If the algorithms required are not standard ones already installed, then ease of upload into multiple TREs is important, such as supporting importing software code or containers from github and docker respectively as previously described. However, if importing and execution needs to be done repeatedly, as an algorithm is iterated for example, support for programmatic workflow management also becomes important. HDR UK will be looking to help coordinate the specification of complex workflow and task execution-oriented specifications such as the GA4GH WES/TES standards that will allow TRE provider to provide orchestrator neutral remote execution functionality.

Best practice for **Federated SAFE Settings:**

FS1. TRE providers should provide ingress and egress (where allowed) to transfer data and code securely between SAFE Settings

FS2. TRE providers should adopt common pseudoidentifier generation processes to enable safe linkage between SAFE Settings

FS3. TRE providers should provide services that allow individuals to remotely execute analysis workflows using TRE supplied tools or externally supplied research software.

UK Health Data Research Alliance

Federation for Safe data

The ability of a TRE ecosystem to support federated queries either through limited programmatic APIs (e.g. Beacons, see federated safe outputs below) or programmatic execution of distributed software packages or containers (see federated safe setting above), depends ultimately on data standardisation. Even more than for Safe Data, adoption of common data standards across the TRE ecosystem is critical. The HDR UK Data Standards Green Paper[43] details appropriate data standards, but only by broadly adopting common data standards will federate data analysis become a practical reality. Initiatives such as by the IMI consortium, European Health Data & Evidence Network (EHDEN)[44] which has been giving grants to support clinical data translation into Observational Medical Outcomes Partnership (OMOP) Common Data Model are leading to growing adoption of this CDM, however there is a long way to go. However, the availability of subsets of data assets transformed into a CDM would still be useful.

Best practice for **Federated SAFE data:**

FD1. TRE providers should consider progressively adopting a Common Data Model, agreed not only across the Alliance but across the wider TRE ecosystem, and developing processes to generate transformations of data assets into this format.

Federation for Safe outputs

There are two main approaches to assessing disclosure risk for output data from TRE – rules-based and principles-based. Rules-based approaches have many implementations and use simple deterministic heuristics (thresholding, rounding, etc) to accept or reject outputs. Some approaches go as far as being able to detect personally identifiable information and obfuscate/reject records. Rule-based approaches tend to be conservative weighing more on preventing disclosure using brute force, rather than considering the utility of the output. The Statistical Disclosure Control Handbook[45] outlines some of the principle-based output checking approaches undertaken by many TRE providers and data custodians. Principle-based output checking evaluations use contextual information about the dataset and project to balance the disclosure risk and utility of the output data. This is a very flexible approach and as such typically undertaken manually and hence takes longer. Both approaches are not mutually exclusive and as such a hybrid approach is typically what is used by TRE Airlock managers.

In terms of harmonising safe output processes across a TRE ecosystem, currently there are no established standards around statistical disclosure control policies and output checking process across TRE providers. Hopefully the creation of a network of Airlock managers and interactions with established organisations

[43] https://www.hdruk.ac.uk/news/hdr-uk-seeks-feedback-on-latest-data-standard-green-paper/

[44] https://www.ehden.eu/

[45] https://securedatagroup.files.wordpress.com/2019/10/sdc-handbook-v1.0.pdf

such as Safe Data Access Professionals (SDAP) working group will encourage convergence of processes and approaches to Safe outputs.

However, there are opportunities for adoption of limited programmatic safe outputs across the TRE ecosystem through the deployment within TREs of Beacons. Beacon software provides externally facing APIs which accept a limited number of query types against data assets held within a safe setting and automatically respond with a statistically safe reply. They are suitable for supporting federated queries and software implementations were originally developed in genomics under the auspices of GA4GH to support distribute anonymous querying across a distributed set of genomics data repositories[46]. Examples of such queries would be to ask each repository if they contained examples of a particular genomic variant, with responses limited to "Yes" or "No" however they can equally be configured to support phenotypic queries. A pilot use case is to enhance the query interface of the HDR Gateway ('Advanced Search') beyond queries limited to the static metadata catalogues provided by Data Custodians for each data asset[47]. The ability to federated queries across different TREs of course depends on the degree of data standardisation across the ecosystem (see safe data above).

Best practice for **Federated SAFE outputs:** (Exemplar from Alliance / Gateway plans)

FO1. TRE providers should consider deploying Beacons within their safe setting to support the programmatic generation of safe outputs from standardised external queries, subject to airlock review processes to approve acceptable query types.

[46] https://www.ga4gh.org/news/extensions-to-the-ga4gh-beacon-api-will-enable-a-more-powerful-community-resource/

[47] https://www.healthdatagateway.org/about/cohort-discovery

UK Health Data Research Alliance

Conclusions and next steps

This paper is an initial step in codifying principles and best practice for implementing and operating Trusted Research Environments, as well as additional principles to construct an ecosystem of TREs. The latter is an area of active work, and examples given are based on existing work to support the Alliance through the Gateway. The Gateway work aims at simplifying and streamlining processes for Data Custodians and TRE providers as well as researchers requesting access and enabling support for federated analysis.

There remains significant work to fully define requirements for implementing and operating TREs, such as frameworks and authorities for accreditation of TREs and guidance on federation, as pilot technical implementations of federation are developed.

The next steps in developing and codifying best practice for TREs and TRE ecosystems will be taken through the workstreams being jointly organised by NHSX and HDRUK for NHS data, as discussed in the forward of this document, and the DARE initiative of UKRI, ADRUK and HDRUK for wider research data. This joint working will provide evidence towards the development of Secure Data Environment policy, which will encompass TREs. These policies for NHS in England will be defined and set by NHSX, part of the Department of Health and Social Care.

Finally, there is considerable potential for this work to have broader impact. Potential ecosystems of TREs stretch wider than the Alliance and wider even than the NHS and contributors to DARE. For example, there are health data projects at a European level being developed that expect to rely on federation between national systems which are in many cases likely to function as TREs. Examples are the **Beyond one million genomes** project (B1MG)[48] which is specifically concerning genome medicine, and the **European Health Data Space**, which is now being planned through a European Joint Action, TEHDAS[49]. Beyond this enabling societal benefits from greater use of data, but in a trusted way, is a major international theme, as recently discussed at G7 meetings during the 2021 UK presidency[50].

[48] https://b1mg-project.eu/

[49] https://tehdas.eu/

[50] https://www.g7uk.org/g7-tech-leaders-agree-bold-new-proposals-to-boost-online-safety-worldwide/

UK Health Data Research Alliance

Appendix A: The case for TREs providing access to health data through safe settings

[This appendix replicates the arguments presented in the TRE Green Paper, with some additional examples]

A central challenge in using health data is how to facilitate research while protecting privacy and so engendering public trust.

The Office of National Statistics (ONS) which facilitates research access to similarly sensitive administrative data described in 2017[15] its role as to *"find a way to maximise the use of the detailed data that ONS holds, while keeping them secure at all times; to let government, academics, businesses and others use these data, while being able to assure you [the public] that you will never be identified, your private details will never become public and that the information you have given us will only ever be used in ways that clearly serve the public good"*. Their approach is summarised as "Five Safes": Safe people; Safe projects; Safe settings; Safe outputs; Safe data. These "Five Safes" were considered as adjustable controls rather than binary settings. Risk is addressed by complementary adjustments on the implementation of each "Safe" to provide an appropriate context for research to occur which maintains an optimal balance between research benefit and overall risk management.

The term "Data Safe Havens" has been used to describe systems for providing researchers with access to data while managing risk of unauthorised re-identification of individuals from de-identified data, however implementations vary considerably[51]. When evaluated against the "Five Safes" defined by ONS, common features of Data Safe Havens are processes of evaluation of research proposals before granting access (Safe projects) and robust processes to de-identify and anonymise data being accessed to reduce the risk of reidentification (Safe data). However, in many cases the model of access is one of "Data Release", i.e. where processed datasets are distributed to researchers, rather than requiring them to carry out their analysis within a controlled environment operated by the Data Safe Haven (Safe setting). Once data is distributed to researchers, controls on who accesses the data (Safe people) and on what is publicly released as results (Safe outputs) and even exactly what research the data is used for (Safe projects) are out of direct control of the Data Custodians. In this model most of the risk-minimisation steps (statistical disclosure control policies) are performed upfront before releasing the data to the researchers, however risks of re-identification of individuals remain. The process also adds complexity both in terms of time and effort for the data custodians to prepare the data, and often reduces the fidelity of the data for research purposes.

Historically organisations such as ONS that have operated safe settings have offered only a limited set of statistical analysis tools to researchers accessing data in their environment. These have been largely adequate given the structure and size of administrative and social science datasets being analysed. By contrast, health datasets, consisting of electronic health records, images, and omics data types, are

[51] Burton, P. R. *et al.* Data Safe Havens in health research and healthcare. *Bioinformatics* **31,** 3241–3248 (2015). https://europepmc.org/article/MED/26112289

typically richer and larger. Analysis is more likely to require complex custom analysis algorithms which correspondingly require more substantial compute resources. Until recently it has been a challenge to provide a safe setting able to support analysis at such a scale and with such diverse tooling.

Recent events and developments have made the provision of safe settings for health data analysis both a desirable and technically practical alternative to data distribution.

Firstly, previous initiatives that have not effectively engaged or consulted on the rationale for data access have resulted in a lack of trust in the phrase 'data sharing'. It has become associated with the risk of jigsaw reidentification through the distribution of data to unknown third parties for unknowable types of analysis and potential unknown further distribution and linkage. All approaches to the de-identification of datasets that contain individual patient level data are limited and require controls enabled by TREs.

Secondly, the adoption of General Data Protection Regulation (GDPR) has made researchers and their organisations subject to serious financial consequences of failing to adequately protect personal health data distributed to them and hosted on computer systems they are responsible for. For many organisations, in particular universities and NHS trusts, it has becoming increasingly challenging to operate computing environments with the required level of security. This is particularly the case when large scale high-performance computing (HPC) is required to support the research community. Similarly, data custodians distributing data have become more risk averse because of potential shared responsibility with receiving organisations for any breach under GDPR.

Thirdly, for large datasets such as images and genomes, data distribution is both inefficient and costly. It results in funders directly or indirectly supporting the costs of storage of multiple copies of large datasets and the associated network costs for multiple large data transfers, when each copy may only be used for a limited period.

Fourthly, over the last 5-10 years the evolution of computer systems have made it practical for researchers to bring complex analysis pipelines to data held on centralised systems[52] and for these systems to be able to support both cost efficient and dynamic scalability of compute[53] for analysis and integral data security[54].

Finally, there are examples of operational systems that provide a safe setting as the only mode of access and that are being used at scale for these classes of health data:

The SAIL DataBank has operated for 12 years with a 'no data leaves' or 'reading library' approach to data access. Using Swansea University's UK Secure e-Research Platform (UKSeRP) to deliver remote access to population-scaled linked data resources from over 400 partner organisations, UKSeRP's fully featured high powered analytical environment has supported hundreds of projects conducted by academics from across

[52] Through Virtual machines or lightweight containers such as Docker.

[53] On premise High Performance Compute (HPC), private or public cloud services

[54] Data encryption at rest under data custodian control via Key Management Infrastructure.

the UK over that time[55]. Similarly, Scotland has implemented a system of federated safe havens with secure analytics platforms, which are safe settings, as described in their Safe Havens Charter[56]. The more recent Genomics England Research Environment (GERE) is also a safe setting and, with >20Pb of genome data from the 100,000 genomes project, operates at a much greater scale. It has >2,000 researchers onboarded to carry out analysis with a range of tools and a full HPC environment (see figure 1).

GERE has the advantages of being set up 1) with explicit research consent from each patient participant; 2) with substantial public and patient engagement and oversight and 3) with completely unequivocal published[57] and public statements that individual level data will not be distributed but will remain within the research environment. As such it has achieved a high level of trust despite dealing with individual genomes, a new and sensitive class of personal health data[58].

UKSeRP has been available as a private cloud offering to third parties wishing to take advantage of UKSeRP's "ready to go" platform to secure and provide access to their own data, under their own governance [59] [60]. As of the 1st January 2020 there were 25 UK-based organisations with UKSeRP tenancies in the UK, including Dementias Platform UK, Avon Longitudinal Study of Parents and Children (ALSPAC), and UK-CRIS (Clinical Record Interactive Search), with an increasing number of installations internationally.

During the pandemic NHS Digital has been able to pilot its TRE service for England[61] for COVID data analysis. Through the Data and Connectivity programme of the COVID-19 National Core Studies[62] it has been part of a pilot to generate federated COVID analysis results across the four UK nations, with UKSeRP, Scotland and Northern Ireland.

Finally most recently OpenSAFELY[63] has shown how a safe setting can be constructed where all analysis operations must be carried out programmatically. While preventing researchers viewing data directly through a VDI makes debugging analysis more challenging, it facilitates complete auditing of all analysis carried out and can enable better transparency and trust. This programmatic approach also facilitates research reproducibility and code reuse.

[55] https://saildatabank.com/wp-content/uploads/SAIL_10_year_anniversary_brochure.pdf

[56] https://www.gov.scot/publications/charter-safe-havens-scotland-handling-unconsented-data-national-health-service-patient-records-support-research-statistics/pages/4/

[57] Turnbull, C. et al. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. BMJ 361, k1687 (2018); Genomics England Protocol https://www.genomicsengland.co.uk/library-and-resources/

[58] https://medconfidential.org/for-patients/loopholes/

[59] https://farrinstitute.org/wp-content/uploads/2018/03/UKSeRP_Case_Study.pdf

[60] https://saildatabank.com/wp-content/uploads/UKSeRP_Brochure_v1.5.pdf

[61] https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england

[62] https://www.hdruk.ac.uk/covid-19/covid-19-national-core-studies/

[63] https://opensafely.org/

It is apparent from patient and public engagement work that there is much greater comfort with data being accessed through a safe setting than data distribution.[64] Where research access is via a safe setting with appropriate patient /public oversight of research activities an opt-out consent model may also be considered sufficient even where data is sensitive, such as the founder CRIS system providing a research environment with Natural Language Processing (NLP) tools on a de-identified copy of the EHR records of the South London and Maudsley NHS Foundation Trust (SLaM)[65]. Even during past media-fuelled public concern about data sharing, the SAIL Databank with its robust proven and robust approach to data curation and access remained uncriticised[66].

Most recently, the OneLondon Citizens' Summit Public deliberation in the use of health and care data identified the following factors that reassured participants in relation to data access[67]:

- Research organisations accessing data within a controlled and secure environment, such as a hospital or research hub, and the data not leaving this environment;
- Access being supervised by appropriate NHS staff or conducted by NHS analysts on behalf of the research organisation;
- Contractual arrangements in place that underpin the data access with consequences for those who break the rules around access (e.g. sharing data outside of the research environment);
- Data not sent or shared outside of the research environment (but could be accessed remotely).

Critical to the success of the TRE-based approach is achieving and maintaining a balance between confidence of data controllers through increased security, benefits to the researcher through improved access to larger datasets, and transparency for public and patients as to who is accessing the data and for what purposes. For researchers it is particularly important that TRE safe settings are well enough supported and designed to ensure researcher can still be productive despite the greater restrictions. Ongoing engagement with all stakeholders is therefore critical, including to communicate the benefits that accompany the move away from the data release model such as the greater security for public and patients as well as improved data access request turnaround times and greater potential for hypothesis-generating or agnostic analysis for researchers.

[64] Great North Care Record 2018 Base: 824 North East representative

[65] e.g. CRIS system, allowing NLP over de-identified mental health records https://www.slam.nhs.uk/research/cris/

[66] Lyons, Ford and Jones, Care.data: why are Scotland and Wales doing it differently? https://www.bmj.com/content/348/bmj.g1702/rr/687637

[67] https://www.onelondon.online/wp-content/uploads/2020/07/Public-deliberation-in-the-use-of-health-and-care-data.pdf

UK Health Data Research Alliance

Appendix B: Timeline of HDR TRE workstream

06/2019	Workshop on TREs at HDR One Institute meeting
09/2019	Established workstream with aim of producing a Green Paper
11/2019	Presentation of early draft Green Paper to HDR external Technical Team
01/2020	Presentation of draft Green Paper to HDR Alliance Board
02/2020	Session on TREs at HDR Alliance Symposium
03/2020	Half day face to face TRE workshop with stakeholders
04/2020	Presentation of draft Green Paper to HDR Public Advisory Board
05/2020	Public consultation on Green Paper launched
06/2020	Online discussion of response to consultation
06/2020	Session on TREs, HDR One Institute meeting
07/2020	HDR Alliance Board approves next steps
07/2020	Updated Green Paper published, with next steps
09/2020	Survey of technical capabilities of existing TREs for National Core Studies
12/2020	Session on TREs, Gateway at UK HDR Alliance Symposium (85 attended)
05/2021	Meeting of TRE working group to discuss planned additions to GP
07/2021	Drafting of TRE Principles and Best Practices

Engagements with Public and other external stakeholders on TREs

09/2019	Presentation to Understanding Patient Data Citizen's Jury on fair partnerships
11/2019	Presentation on TRE implementation to Pistoria Alliance
02/2020	Session on TREs at OneLondon Citizen's Summit
06/2020	Presentation on TRE Strategy at NHS-HE Forum
10/2020	Talk on TREs at Westminster Health Policy Forum
11/2020	TRE demo and Teach in for Office of Life Science (from Genomics England)
11/2020	My data Event: Introduction to TREs [with Q/A followup 12/2020]
02/2021	Public Policy Projects talk, including TREs

Contact
enquires@hdruk.ac.uk

DOI: https://doi.org/10.5281/zenodo.5767586

Visit the UK Health Data Alliance website: https://ukhealthdata.org/