

GP Data for Planning and Research
Check and Challenge Group – TRE Showcase
Minutes of the meeting held 9th November 2022
Final – Approved for publication

Attendees - Members

Name	Organisation	Name	Organisation
Helen Stokes-Lampard Chair	Academy of Medical Royal Colleges	Ruth Ellenby	Royal College of General Practitioners
Ian Hulme	Information Commissioners Office	Paul Atkinson	Royal College of General Practitioners
Tim Donohoe	NHS Transformation Directorate	David Parkin	British Medical Association
Nicola Byrne	National Data Guardian	Mark Coley	British Medical Association
Nicola Perrin	Association of Medical Research Charities		
NHS Digital Programme Team Representatives			

Attendees – Presenters

Name	Organisation	Name	Organisation
Ben Goldacre	Nuffield Department for Primary Care Health Sciences (OpenSAFELY)	Sebastian Bacon	Nuffield Department for Primary Care Health Sciences (OpenSAFELY)
Amir Mehrkar	Nuffield Department for Primary Care Health Sciences (OpenSAFELY)	Ian Diamond	Office for National Statistics
Michael Chapman	NHS Digital	Alison Pritchard	Office for National Statistics
Elizabeth Way	NHS Digital	Bill South	Office for National Statistics

Attendees - Observers

Organisation
NHS Digital x 8
OpenSAFELY x 1
Academy of Medical Royal Colleges x 1

Noted Declarations of Interest

Declarations noted against each presentation

OpenSAFELY Presentation

Ben Goldacre

- Speaking today as Director of the Bennett Institute at the University of Oxford and PI of OpenSAFELY

Additional DOI

I am a clinical academic, Bennett Professor of Evidence Based Medicine and Director of the Bennett Institute for Applied Data Science at the University of Oxford: bennett.ox.ac.uk. Our research covers a range of areas in epidemiology and data science, a full list of current papers is at <https://www.phc.ox.ac.uk/team/ben-goldacre>. Major projects include OpenPrescribing.net (an open access data analysis tool for prescribing data), OpenSAFELY.org (a portable open-source software platform for secure analysis of data), OpenCodeLists.org (a tool for curating patient data) and TrialsTracker.net (an audit tool for clinical trial reporting), TrialsTracker.net (an audit tool for clinical trial results reporting). All outputs from the Bennett Institute, including all analyses and tools, are shared as open source and open code free for review and re-use by all. My group at the University of Oxford has received funding via the University of Oxford from a wide range of public and charitable funders: the NHS National Institute for Health Research (NIHR), NHS England, the NIHR School of Primary Care Research, the NIHR Oxford Biomedical Research Centre, the Laura and John Arnold Foundation, the Mohn-Westlake Foundation, NIHR Applied Research Collaboration Oxford and Thames Valley, UKRI / MRC, the Wellcome Trust, the Good Thinking Foundation, Health Data Research UK (HDRUK), the Health Foundation, and the World Health Organisation. My group at the University of Oxford continues to seek funding and collaboration from these and similar organisations. I receive personal income from speaking and writing books (and, infrequently in recent years, journalism) for lay audiences on the use and misuse of science and evidence. In policy work: I chaired the HealthTech Advisory Board for Sec of State; I have sat on various local and national committees; I am a Non-Executive Director at NHS Digital; I led the Goldacre Review ("Better, broader, safer: using health data for research and analysis" March 2022) for Secretary of State for Health and Social Care. I am an advocate of secure approaches to managing health data; and open code and "reproducible analytical pipelines" as a technical means to support modern, open, collaborative, productive and high-quality work.

Sebastian Bacon

- Speaking today as CTO at the Bennett Institute and responsible for the OpenSAFELY software

Amir Mehrkar

- Represents the RCGP / BMA in PAG for advice to NHSD about access to GDPPR data for Covid-19 research. And in that role, OpenSAFELY (my main role) has been identified as a TRE that is supported for GP Data processing by the profession. Similarly, ONS and NHSD TREs have been identified as appropriate

ONS Presentation

Prof. Sir Ian Diamond

- Speaking as the National Statistician

Alison Pritchard

- Speaking as the Deputy National Statistician and Director General for Data Capability at the Office for National Statistics (ONS), and SRO for the Integrated Data Service

Bill South

- Speaking as Acting Deputy Director, Research Services and Data Access division

NHS Digital Presentation

Michael Chapman

- Speaking as NHS Digital Director of Research and Clinical Trials and primary director responsible for the NHS Digital Secure Data Environment

Additional DOI

I'm an independent member of the UKSA Research Accreditation Panel in addition to my NHS Digital role and as such, I am part of the group that accredits the ONS Secure Research Service.

Elizabeth Way

- Speaking as Product Manager of the NHS Digital Secure Data Environment and fully employed by NHS Digital

Meeting Notes

Meeting Notes

Introductions, agenda, and declarations of interest

- Chair welcomed members to the call and outlined the focus for this session was to see presentations and demonstrations of the three TREs noted in the 2021 ministerial letter
- The focus of the session is to build member awareness of the three systems to better support evaluation that the NHS Digital will meet the ministerial commitments prior to a future go live event
- Chair noted that given the broader focus of the session observers have been welcomed and asked any speakers to clearly introduce themselves for the benefits of all attendees
- Chair noted that given the nature of the presentations all attendees should also declare any conflicts of interest prior to directly participating in the session

OpenSAFELY presentation

- Chair passed to OpenSAFELY presentation team and noted declaration of interests

- Key presentation points noted were:
 - The data currently used within OpenSAFELY
 - Volumes and examples of the research outputs to date
 - How OpenSAFELY compares to traditional research data access models
 - Key challenges within the traditional models that OpenSAFELY is designed to address
 - Noted that while OpenSAFELY is currently focused on TPP and EMIS deployments, OpenSAFELY is fundamental a set of open-source tools and a team that can be deployed in any data centre to make a TRE
 - An overview of how OpenSAFELY works including the federated model across the underlying data processors, and code sharing on GitHub
 - The current user base and how access requests are managed
 - Key user support features such as the co-pilot service

- Following the initial presentation, a walkthrough of the user journey and support metrics was presented. Key points noted were:
 - Public facing audit records
 - Back-office management processes and metrics
 - Development update for OpenSAFELY interactive – a point and click service to support users without the need to code their own queries

- Chair thanked team for the presentation and opened up to members for questions. Noted questions were:
 - How useable is it for analysts given the programming languages used?
 - The use of Python and R is not proving a barrier to useability at present. The support documentation and co-pilot service provide users with a foundation to use the service, and the development of the point and click toolset will make it even easier
 - What is the volume of access requests?
 - The initial volume of requests was significant and exceeded the capacity and funding of the support team to address. This is being worked through, but new users are not being accepted at present
 - How many outputs are rejected by the manual output checking?
 - An audit has not been completed of the rejection rate of outputs and it would be good to undertake this to provide visibility. The challenge remains the volume of outputs for the team to get through rather than high rejection rates

Office for National Statistics (ONS) presentation

- Chair passed to ONS presentation team and noted declaration of interests

- Presenters noted the presentation will be split between the established Secure Research Service (SRS) and the new Integrated Data Service (IDS), which is in development at present.

- Key SRS presentation points noted were:

- Overview of the current service user base, live project volumes, and data sets
- Overview of the three elements of the SRS
 - SRS Metadata Catalogue – Supports initial concepting
 - Research Accreditation Service – Ensure users and projects are accredited
 - SRS Platform – Where accredited projects run their projects
- Overview of how the SRS fully integrates the 5 Safes model
 - User access methods determined by Safe Settings approach
 - Full user journey presented from 5 Safes perspective
- Key IDS presentation points noted were:
 - SRS service is performing well but reaching limits of its capability
 - IDS service intends to address capability challenges of SRS
 - Premise is to link across disparate datasets through a shared indexing mechanism
 - IDS is current in public beta stage
 - Overview of the principles on which the IDS is being designed
 - High-level explanation of how the IDS will work and how it differs from the SRS
- Chair thanked team for the presentation and opened up to members for questions.
Noted questions were:
 - Who will oversee the data linkages within the IDS to ensure they do not undermine data security?
 - Plan to manage within ONS with all new linkages quality assured and a set of pre-approved linkages available immediately for users. Potential exists for process for users to manage indexing process for new linkages, but this will need to be developed
 - Who can use these services and what skills are needed?
 - Focus is on making the services as useable as possible by using the current most popular programming language (Python) as well as others and providing support for users. Expectation is majority of data researchers would be able to use this quickly.
 - The system is optimised for quick use and not deep technical knowledge. The purpose was to provide answers to high-level and ministerial questions as quickly as possible, so speed is key

NHS Digital presentation

- Chair passed to NHS Digital presentation team and noted declaration of interests
- Presenters noted the presentation will be given as a live demonstration of the service preceded by a short introduction.
- Key points from introduction were:
 - Focus on revamping technology rapidly stood up during the pandemic so that it is suitable for long term future use
 - NHS starting from a position where dissemination is standard to moving to a data access service

- Focus is on matching the service to the specific use case to ensure what data is accessed and the level of access is appropriate for each situation
- Positioned as a journey to take the system from where it is now to a service comparable to the other services demonstrated today and that can become the default route for external users to access data held by NHS Digital.
- Current terminology being used is a Secure Data Environment (SDE) rather than Trusted Research Environment (TRE), as per recent NHS publications
- Key points from live demonstration were:
 - Overview of user login narrated (done before demonstration for efficiency)
 - Service uses off-the-shelf components brought together by NHS Digital team
 - Access permissions managed before access and SDE only allows access to data that has been approved for that user
 - Based on 5 Safes model and all processes map back to elements of the model
 - A choice of programming languages is provided for users
 - Overall access and user environment is set based on the agreed data sharing agreement (DSA) for each approved project
 - Projects are shared across all users who are on the DSA
 - Has functionality to re-use code from your history or from other users within your project/DSA
 - Code can't be currently shared outside of each DSA. This is under review, but focus is on finding a safe and secure process to do so
 - Admin function demonstrated showing how access is limited based on the DSA
 - Audit logs are collected in real time and capture all required detail
- Chair thanked team for the presentation and opened up to members for questions. Noted questions were:
 - Does this service impact the requirement for GDPR?
 - GDPR will still be required to bring the data into the SDE, and the SDE will be the secure environment which controls access and usage
 - Will these solutions drive more network traffic (particularly IDS)?
 - ONS team responded that the IDS may drive more traffic due to the linkages across datasets and they are working with partners to ensure an appropriate solution, but this will depend on the specific analysis being done and what linkages are required

Conclusions and AOB

- Members queried what the expected outputs and next steps were from this showcase
 - Session was aimed to raise awareness of the ONS and OpenSAFELY systems given the specific reference within the ministerial commitments letter that these services represented best practice for the NHS Digital service to design against
 - By building understanding within the group's members of these services, the focus is to allow members to better support the GDPR programme team to ensure the NHS Digital service meets the ministerial commitments

- It was noted that there is an opportunity for members of this group to use this platform to answer some of the big questions facing the use of data for public good, but this was agreed as outside of the scope of this particular advisory group
 - It was agreed that the key next step for the programme team was to brief the new ministerial team to revalidate the forward approach for GDPR and the ministerial commitments. This was expected to be accomplished within November to allow discussions to progress
 - Check & Challenge members will be kept informed of all progress in this area, and the group will be reconvened following ministerial input
- Presenters and members were thanked for their time and the call was ended by the chair