

OpenSAFELY Software Oversight Group Agenda

[04/06/2025]

ATTENDEES

Chair:

- Prof. Sir Nigel Shadbolt

Bennett Institute:

- Prof. Ben Goldacre - Director and OpenSAFELY PI
- Dr. Amir Mehrkar - Head of IG and external relationships
- Seb Bacon - Chief Technology Officer
- Pete Stokes - Director of Platform Development
- Freddie Longfoot, PPIE Lead

External Group Members:

- Tariq Khokhar - Head of Data for Science & Health, Wellcome Trust
- Jeni Tennison - Open Data Institute
- Dr. Alexandra Freeman - Crossbench peer, House of Lords
- Prof. Stephen Evans - Prof. of Pharmacoepidemiology, LSHTM
- Michael Chapman Director of Data Access & Partnerships, NHS England
- Dr. Imran Khan - RCGP
- Dr. Mark Coley - BMA
- Sam Smith - MedConfidential
- Dr. Chris Bates - Director of Research and Analytics, TPP
- Expert patient representative from Digital Critical Friends - Kevin Minier
- Calum Davey - Executive Director of Research and Best practice, NIoT
- Prof. David Clark - Prof of Experimental Psychology, University of Oxford

Invited:

- Nin Sandhu, Head of Innovation & Collaboration, NHSE

Secretariat:

- Jade Womersley, Executive Assistant, Bennett Institute

Apologies:

- Tariq Khokhar - Head of Data for Science & Health, Wellcome Trust
- Dr. Shaun O'Hanlon - Chief Medical Officer, EMIS Group
- Rony Arafin - Chief Operating Officer, AphA CIC
- Prof. Laurie Tomlinson - Prof. of Clinical Epidemiology, LSHTM and Expert OpenSAFELY Researcher

AGENDA

Item	Action	Owner	Status notes
1	Introduction from the chair <ul style="list-style-type: none"> Request for AOB - Ratify previous minutes - ratified and added to the OpenSAFELY Governance page. Review of actions 	Nigel	10:00 - 10:05
2	Team Update <ul style="list-style-type: none"> NHSE OpenSAFELY Service Staffing Funding Outputs Tech Patient and Public Involvement and Engagement (PPIE) OpenSAFELY Schools 	Pete	10:05 - 10:30
3	NHSE Update <ul style="list-style-type: none"> Legal basis for Non-COVID research Commercial Arrangements with General Practice System Supplier (GPSS) 	NHSE	10:30 - 10:45
4	Planning for non-COVID	Pete / Michael Chapman	10:45 - 10:55
5	Break		10:55 - 11:05
6	NHS Talking Therapies (TT)	Seb	11:05 - 11:20
7	Global Burden of Disease (GBD)	Ben	11:20 - 11:35
8	Engagement with National Data Library (NDL) / Health Data Research Service (HDRS)	Ben	11:35 - 11:50
9	AOB		11:50 - 12:00

Minutes

1. Introduction from the Chair

- **Actions:**
 - Request for any other business (AOB)
 - Ratification of previous meeting minutes now published on the [OpenSAFELY Governance page](#)
 - Review of previous actions – all confirmed as completed
-

2. Team Update

- **NHSE OpenSAFELY Service:**
 - OpenSAFELY continues to support COVID-19 research; work on new legal Direction to expand service remit to enable non-COVID research is in its final stages. Market size for COVID related work is now small.
- **Staffing:**
 - High volume and quality of applicants for all OpenSAFELY roles
 - Tech team focus is on maintaining the service, and then three priority themes:
 1. New Deployments in new data centres
 2. Data Curation
 3. Dashboards
 - A visual organogram is being developed to illustrate the current structure and will be shared at the next meeting
- **Funding:**
 - Bennett Institute (BI) in a reasonably stable financial position for 18 months due to NHS England (NHSE) contracts and Wellcome Trust grants.
 - Mentioning two new research (not infrastructure) grants for Early Meet and Mid Career Researchers (EMCRs) on team, NIHR grants, useful to have strong epidemiology teams from our organisation and others associated with infrastructure teams to maintain strong delivery of tools.
 1. Respiratory vaccine evaluation (starts Oct 2025).
 2. Heart failure study (starts June 2025)
- **Outputs:**
 - Continued delivery of high-impact research and tools/ two preprints and one published paper since January.
 1. Fewer publications to highlight than in previous meetings, due to diminishing interest in research related to COVID-19. All are very keen that NHSE completes the non-COVID permissions work so that the community can make good use of the GP data that has been made available by the OpenSAFELY team.
- **Tech:**
 - Enhancements to tools and data (e.g. frailty index, discharge data); changes and enhancements to R support; preparing for non-COVID expansion
- **Patient and Public Involvement and Engagement (PPIE):**
 - Online workshops completed; in-person sessions being planned.
 - Collaboration with NHSE Comms on messaging
 - Patient Advisory Group recruitment for Talking Therapies begins July 2025

- **OpenSAFELY Schools:**

- The BI is helping the National Institute of Teaching (NlOT) prepare for their annual review with funder XTX Markets by developing an OpenSAFELY Schools prototype, which has received positive feedback and aims to secure further funding
 - Various approaches are being explored for use of Electronic Health Research Query Language (ehrQL) with educational data that has many structural similarities with Electronic Health Record (EHR) data and other event-list or “transaction” data.
-

3. NHSE Update

- Legal Basis for Non-COVID Research:
 - New directions largely approved by relevant internal and external stakeholders.
 - Type one opt-out already respected within OpenSAFELY
 - Working with BI and stakeholders to ensure that National Data Opt-Out can also be respected, where requested or required for specific projects
 - Commercial Arrangements with GPSS:
 - Delays in contract between NHSE and Optum continue, but progressing.
 - Optum has systems ready for Data Protection Notice (DPN) comms but contractual negotiations are still underway
 - Two options discussed:
 - Launch in stages (TPP first, then Optum)
 - Wait for simultaneous launch (risk of delay)
 - A discussion was raised by members about the significant concern in the wider community about whether engagement and permissions were done correctly on the Health Data Research UK (HDRUK) led “AI Foresight” project that was conducted in another Trusted Research Environment (TRE) (not OpenSAFELY); and the potential for public, patient and professional confusion - specifically that some may incorrectly think this was an OpenSAFELY project, and whether this is a reputational risk that could spread to harm OpenSAFELY.
 - A request was made from the Board to NHSE to ensure that there is clear reactive and proactive communication throughout to ensure that it is understood that OpenSAFELY is not related to Foresight in any way.
 - Safeguards in OS highlighted: transparency, public codebase, GP data control
-

4. Planning for Non-COVID Research

Various aspects of work discussed including:

- Technical Work:
 - To ensure that is only accessible for all practices that acknowledge Direction has been issued
- Information Governance (IG):
 - New Data Access Agreements in progress
- Service Development:

- BI/NHSE collaboration to create a process consistent with existing NHSE DARS process; feedback that it is important to ensure an application process that is swift, but also selects high value projects that demonstrate value and impact from the platform and the data.
- Policy:
 - Updates to reflect new directions and any changes in how OpenSAFELY operates

6. NHS Talking Therapies for Anxiety and Depression (TTad)

- Project Goal:
 - Link TTad data with GP data via OpenSAFELY and make it accessible to a wide community of users for high impact research and analysis.
- Funding:
 - Includes support for a new kind of Electronic Health Record service: ehrSU (Support Service)
- Data & Research Potential:
 - 98% of TTad users provide outcome data
 - OpenSAFELY enables large-scale linkage (8.5 million patients), improving long-term outcome tracking
 - Addresses limitations of small sample sizes and DARS access barriers
- Policy & Support:
 - Strong backing from Department of Health and Social Care (DHSC) and clinical mental health community
 - Potential expansion into community and children's mental health
- PPIE:
 - New experience-led additional public advisory group being formed
 - Focus on governance, consent, security, and acceptable data use
 - Potential for co-development with Digital Critical Friends (DCF)
- Tech
 - Tech and policy discussions between BI and NHSE on innovative and even more intensive data minimisation work as the best model to deliver access while maximising privacy and minimising data flows; the "Take Only What You Need" model.
- Timeline:
 - Legal basis → dataset integration → minimised data transfer and linkage

7. Global Burden of Disease (GBD) Collaboration

- Overview:
 - GBD is regularly described as the world's largest health research project, producing global disease statistics in comparable forms across a huge number of diverse nations.
 - Detailed discussions with GBD team including three in-person visits from them to OpenSAFELY team, on deep collaboration for federated data infrastructure.
 - GBD is interested in OpenSAFELY's privacy-preserving architecture, and the data management benefits of ehrQL.
 - Substantial strategic value: additional demonstration of OpenSAFELY use as a general-purpose federated data analysis service with data remaining in situ.
-

8. Engagement with NDL / HDRS

- NDL (National Data Library):
 - Original vision: network of standalone nodes for data preparation.
 - Set out in paper: <https://zenodo.org/records/14671714>
 - Very substantial engagement by OpenSAFELY teams with NDL team including their Senior Leadership Team (SLT).
 - Very substantial turnover in NDL SLT since then.
- HDRS (Health Data Research Service):
 - Strategic direction not yet clear
 - OpenSAFELY always keen to be involved, this is anticipated, under discussion

9. AOB (Any Other Business)

- AI and Data Infrastructure:
 - Discussion point raised by a member who had heard other organisations state that “OpenSAFELY can’t do AI” - others had heard similar, but these claims were typically non-specific and conveyed second-hand, so difficult to respond to or refute.
 - The reported claims are clearly incorrect: OS can support AI, as it is a wrapper around compute (so, for example, Large Language Models (LLMs) could run), but this would require discussion with, and explicit support of, stakeholders (inc. this group) to ensure a situation like that surrounding “AI-Foresight” is avoided.
 - Various options to respond to the claims were discussed, including proactive comms and work to demonstrate that AI models work in OpenSAFELY. The group agreed it was likely sufficient, at the current time, to simply correct these claims when made, as experts in the space will already recognise that they are incorrect.
 - Team has various design ideas for a GP EHR TRE specifically tailored to AI uses, were that to be something that UK funders were able to resource development of.

Item (Year- Meeting- Action)	Action	Owner	Status Notes
2025-02-01	Name of group meeting changed from OpenSAFELY Oversight Group to OpenSAFELY Software Oversight Group	N/A	Completed and next meeting will be renamed
2025-02-02	Bennett Institute Organogram to be agreed and circulated	Jade/ Klaudia	Being finalised and will be shared at next meeting

Glossary

DARS - The Data Access Request Service (DARS) provided by NHS England *is the official system used to apply for access to NHS health data for research, planning and other public interest purposes.*

Data Access Agreement - is the agreement between the data provider and researcher that explains the terms and conditions for accessing the data.

Data Curation - is the organisation and cleaning up of healthcare information for the benefit of doctors and researchers, ensuring that information is accurate, consistent and usable for patient care, research, and analysis.

Data Protection Notice - is a document which explains how an organisation collects, uses, stores, and protects your personal data.

Digital Critical Friends - is an independent public advisory group set up to provide constructive feedback to improve processes or technologies, and act as a critical voice to the OpenSAFELY platform developments.

Federated Data Infrastructure - is a system which allows multiple organisations to collaborate and share data without it needing to move *from its original location.*

NIHR grants - NIHR grants provide funding and support for health and care researchers, to develop skills and carry out independent research that seeks to improve patient and care outcomes.

R - is a programming language and software environment designed for statistical computing, data analysis, and visualisation.

Talking Therapies - (also called psychological therapies) are effective and confidential treatments delivered by fully trained and accredited NHS practitioners to help people struggling with feelings of depression, excessive worry, social anxiety or post-traumatic stress disorder (PTSD).

Trusted Research Environment - is a *highly secure digital platform* where researchers can access *and analyse* sensitive personal data safely, without compromising privacy or confidentiality.

Type one opt out - is the process which enables patients to opt out of their data being used for research, planning services and public health analysis, permitting the sharing of their data for their direct care only.