This protocol has regard for the HRA guidance and order of content for study protocols

Prepared with reference to Section 11 (Research Databases) of the Standard Operating Procedures for Research Ethics Committees Version 7.6 Sept 2022

FULL/LONG TITLE OF THE STUDY: THE SMARTPHONE AND WEARABLE DATA FOR HEALTH RESEARCH DATABASE

SHORT STUDY TITLE / ACRONYM: SMART-Health

PROTOCOL VERSION NUMBER AND DATE: Version 1.0 4/04/2025

Sponsor/Data Custodian: University of Sheffield

Chief Investigator: Tim Chico

Study Governance Administrator: Tim Chico

Telephone:

Email: t.j.chico@sheffield.ac.uk

RESEARCH REFERENCE NUMBERS

University of Sheffield	191893
IRAS Number:	337916
Sheffield Teaching Hospitals Registration Number	STH23499

KEY STUDY CONTACTS

Chief Investigator	Tim Chico
Co-Investigators	Vita Lanfranchi Ellen Buckley Suzanne Mason James Wild Chris McDermott Steve Haake George Peat

	Márjory da Costa Abreu Richard Clayton Sanja Dogramadzi Robert Storey Sheeba Koilpillai Nick Hamilton Rachel Cliffe Hannah Clemmens
Sponsor	University of Sheffield Western Bank Sheffield S10 2TN

Table of Contents

iii. STUDY SUMMARY

Full Title	THE SMARTPHONE and WEARABLE DATA FOR HEALTH RESEARCH DATABASE
Short Title	SMART-Health
Study design	Observational Consented Research Database
Study Participants	Adults able to provide informed consent, who own or are able to use the required smartphone and/or wearable devices, and who use email.
Planned Sample Size	10,000
Planned Recruitment Period	10 years (planned to be renewed every 5yrs)
Follow up and data collection duration	Until participant withdrawal, study closure, or death of the participant
Study purpose	This database will be used for the following types of observational research: ● Discovery of generalizable, interoperable measures of health from smartphone and wearable data ● Discovery of diagnostic signatures of smartphone/wearable data ● Discovery of predictive signatures of smartphone/wearable data ● Discovery of strategies to monitor people with specific diseases for signs of response to treatment, adverse outcomes or complications
Description of Database	To establish a research database of smartphone and wearable data linked with the participant's routinely collected health-related data from the NHS and other data sources. Data will be held within a secure data environment. Researchers will apply for access to the linked, de-identified data via a SMART-Health Data Access Committee. A generic research database ethics approval is sought to allow the Data Access Committee to approve appropriate projects from bona fide researchers that seek to examine how smartphone and/or wearable data could be used to predict, prevent, diagnose or monitor human diseases.

iv. KEY WORDS

Smartphone data, wearable data, disease prediction, diagnosis, disease monitoring, digital technologies

1 Background & Rationale

Why is smartphone and wearable data useful for health research?

Smartphones and consumer wearables (like smartwatches) collect a wide range of data relevant to health. This includes measures of mobility, physical activity, and physiology (such as sleep, heart rate, oxygen level). These devices also allow the user to respond to questions about symptoms, quality of life, mood, etc. Collecting this data over months or years provides insight into how health changes over time.

In the UK over 90% of people use a smartphone and 40% use a wearable device such as a Fitbit, Apple watch or Garmin smartwatch[1]. However, despite the high rates of use in the population, **researchers have little access to useful repositories of smartphone and wearable data.**

Even large-scale, well-funded UK research databases such as UK Biobank[2] and Our Future Health[3] do not collect smartphone and wearable data from consumer devices.

The US study All of Us has recruited 16,000 Fitbit users who provide access to their wearable data but no smartphone data[4]. Smartphone data is more likely than wearable data to provide insights that are generalisable as wearables are used by more affluent participants and so the data generated is less representative of the wider population.

UK Biobank does collect measures of physical activity and ECG from “research-grade” wearable sensors, providing these to selected participants for a week only[5]. This allows standardised data from single devices and allows any consenting participant to provide data. However, it prevents long term data capture and applicability to healthcare, as such devices are not used clinically or by consumers.

An approach that uses the person’s own devices (called Bring Your Own Data, BYOD(Dixon et al. 2023)) provides data over much longer periods than when devices are provided and returned, and insights obtained are more likely to lead to useful clinical tools that could be introduced to the entire population of millions of device users.

Smartphone and wearable data need to be linked with other health-related data

Smartphone and wearable manufacturers already hold large amounts of smartphone and wearable data on users, with consent to collect and analyse data included in the terms and conditions (which users rarely read). This varies across manufacturers: Apple does not hold data, which is stored on the device only.

Despite having access to large amounts of smartphone and wearable data, manufacturers alone cannot perform the research required to understand how to use this data in healthcare. For this purpose, smartphone and wearable data must be “linked” with other data about the user and their health(Sudlow 2024).

At the most basic level, age, location, gender, and other demographics are necessary to understand differences in the patterns of data between these groups.

However, for most research, smartphone and wearable data also needs to be linked with data related to the person’s past, present and future health (such as if they have ever had cancer, are currently taking a particular medication, or if they develop dementia in the future). This health data is essential for research to establish the associations between smartphone and wearable data and health.

Until recently, an individual's health data (such as NHS record) was predominantly kept in paper records that were very difficult to access for research. The increasing digitisation of healthcare data via electronic patient records (EPR) and nationally held datasets provides the opportunity to link a person's smartphone and wearable data with this electronic medical data. With appropriate approvals, governance and safeguards, clinical data from NHS records and other health-related data (such as housing and social care) held by organisations, including primary care, secondary care, and national bodies such as NHS England can be linked with smartphone and wearable data.

What is best practice for privacy protection and data security?

The COVID pandemic, increasing digitisation in the NHS, and improved technical capabilities for data storage and analysis all greatly accelerated use of routinely-collected NHS data for research. Because NHS data is primarily collected for healthcare purposes, not for research, consent for "secondary" research on unconsented data has usually not been obtained (although there is an ability to opt-out). Approval for such research is sought via the Confidentiality Advisory Group (CAG).

To protect the privacy and security of routinely collected NHS data used for research without consent, such data is increasingly held in "Secure Data Environments" (SDEs). Instead of providing copies of data to researchers (which could be shared with unapproved persons or used for non-approved purposes), approved researchers access the data within a secure environment by connecting via an authentication process, and only access de-identified data required for their project. The NHS England Data for Research and Development programme has established "Subnational Secure Data Environments" tasked with bringing together all NHS data for this purpose.

Analyses (aggregation, selection of groups of individuals for analysis, statistical comparisons etc) are all performed within the SDE's computational and storage environment, not on the personal computer of the researcher. Approaches can be applied to protect privacy (such as low-number suppression where a set of filters that would return a very small number of individuals (and possibly allow re-identification) are rejected.

Only anonymous aggregated analyses can be exported (and such outputs are checked to ensure no identifiable data is exported from the secure environment). Although this approach was developed for research on non-consented NHS data, it is increasingly being used for consented research data (such as by UK Biobank). This applies the highest standards of security to consented data and is particularly appropriate when large amounts of potentially sensitive data are to be collected over long periods of time.

NHS and other health-related data exist in multiple locations with different data controllers and access procedures. For example, hospital episode statistics (HES) data from England that details inpatient diagnoses is held nationally and can be accessed from NHS England via a single application, while primary care data is held by multiple primary care organisations, requiring multiple data sharing agreements. A similar situation exists for lab test and imaging data which is held by large numbers of different secondary care organisations across the UK.

The 2024 Sudlow Review[6] highlighted the importance of bringing all NHS data together and making it available for research with appropriate safeguards including the need to link it to smartphone and wearable data (with consent).

The University of Sheffield has established a SDE in South Yorkshire called **Data Connect**. Data Connect is part of the Yorkshire and Humber NHS Subnational Secure Data Environment. In

partnership with the South Yorkshire and Bassetlaw Integrated Care Board (the overarching NHS organisation for the region), Data Connect is bringing together routinely collected de-identified NHS data for research.

The need for SMART-Health

There is a pressing need for a large-scale research database of Smartphone and Wearable Data linked with participants' NHS and other health-related data for bona fide researchers to conduct approved research that aims to understand how smartphone and wearable data might be used to improve human health.

Such a database would allow a wide range of research across almost all diseases. For example:

- How patterns of smartphone and wearable datadiffer in people who later develop specific diseases, compared with those who remain healthy. This could lead to **new ways to predict and prevent** such diseases.
- How smartphone and wearable data differs between healthy people and people with different diseases. This could identify **potential diagnostic and treatment approaches**.
- How smartphone and wearable data changes over time in different diseases and with ageing. This could identify **ways to monitor and prevent deterioration**.

This protocol describes the establishment of such a research database, called **SMART-Health**.

SMART-Health will be held within a secure data environment (SDE). We will obtain smartphone and wearable data from up to 10,000 consenting participants and link this to the individual's routinely-collected NHS data.

Justification of recruitment target

SMART-Health is designed to support a wide range of specific research projects. The sample size and statistical power for each specific project will vary.

Previous work from a similar research database All of Us shows that high-impact research on the links between Fitbit data and future disease can be performed with acceptable statistical power with under 7,000 participants.

Therefore SMART-Health will aim to recruit 10,000 participants in its first phase (5yrs). We expect to expand recruitment in future phases. A previous study within our group using the Active10 application registered 307,834 users from Great Britain in 2yrs therefore we believe this is a reasonable target.

Data Management and Access for Research

To protect participants' privacy, all data will be de-identified by the core team before researchers have access and before research is performed. This de-identified data will be held and accessed via a secure data environment (SDE) and will not be freely shared.

A local Data Access Committee will evaluate, approve or reject applications to access this data. The scope of this research must be to understand the links between smartphone and wearable data and past, present or future disease (which may lead to ways to predict, prevent, diagnose or monitor disease). This approval process will follow the "Five Safes"(UK Data Service 2021) (Safe Data, Safe Projects, Safe People, Safe Settings, Safe Outputs).

2 Non-technical (Plain English) Summary

Over 90% of UK adults use smartphones, and almost half use wearable devices like smartwatches.

Data from these devices (such as heart rate, steps per day, or hours of sleep per night) might help researchers discover new ways to predict, diagnose and monitor different types of disease.

To make these breakthroughs, data from the smartphone and wearable needs to be linked with health-related data such as from the person's medical record.

We will recruit 10,000 participants who give permission for us to collect smartphone and wearable data from such devices. Initially this data will measure physical activity, quality of life, sleep and heart rate. We will also ask if participants want to opt-in to sharing their location data.

All smartphone and wearable data for research will be "de-identified" by removing features that would allow a person's identity to be revealed. This means the identities of the people providing data will not be revealed to the researchers using the data or in the research findings.

We will link this de-identified smartphone and wearable data with the person's health-related data within a research database held in a secure, safe environment at the University of Sheffield.

Researchers will apply to a SMART-Health Data Access Committee (that includes public members) to access specific types of data within the secure data environment.

The committee will approve access to data for a wide range of research projects in the public good that aim to improve prediction, prevention, diagnosis or monitoring of disease (such as cancer, heart disease or other major health problems) while ensuring the privacy of the people providing the data.

3 Aim and Objectives

3.1 Study Design

SMART-Health is a prospective, longitudinal, observational cohort of UK adults with and without disease leading to establishment of a research database for use in a wide range of health research projects.

3.2 Aim

The aim of SMART-Health is to allow researchers to explore the insights that smartphone and/or wearables can provide when linked with participants' routinely collected NHS data to enable a wide range of research into prediction, prevention, diagnosis and monitoring of disease. Our objectives are to:

- Establish best practice mechanisms to approach, inform, and recruit the widest possible range of consenting participants with and without different types of health conditions (target 10,000 people total)
- Obtain consent to allow access to participant's smartphone and wearable data (personal and/or provided devices) and link this with their routinely-collected health-related data (particularly NHS data) for research
- Enable recruitment of participants in current or future research studies to allow additional insights from smartphone and wearable data to these studies.
- Obtain longitudinal smartphone and wearable data from participants, alongside demographic information, and user generated data
- Securely store and link de-identified smartphone and wearable data with de-identified routinely collected NHS and health related data within a Secure Data Environment

- Allow bona fide researchers performing projects within scope to access project-specific data extracts within a secure data environment.

4 Methodology

4.1 Project plan

The SMART-Health research database will collect smartphone and wearable data (detailed in section 5) from consenting participants with different health conditions and characteristics and link this with their NHS and other health-related data.

To enable large-scale recruitment, and maximise scientific value, SMART-Health will:

- Recruit a diverse range of participants from the general population and patient populations.
- Enable remote recruitment and consent so that face-to-face interaction with a researcher is possible but not essential.
- Obtain specific types of smartphone and wearable data relevant to health (section 5) from the participant's own devices
- Where appropriate and as funding allows, provide devices to participants who do not currently have these to allow them to take part
- Link de-identified smartphone and wearable data with demographic data and routinely collected NHS data from the South Yorkshire Integrated Care Board and other data custodians including primary care, secondary care and national databases within a secure data environment.
- Provide approved researchers access to this linked data for specific approved research projects via a Data Access Committee.

4.2 Inclusion/Exclusion Criteria

All SMART-health participants must be:

- Over 18 years old at the point of recruitment
- Able to understand the participant information sheet and give informed consent
- Current user of a smartphone or is willing to use one
- Has an email address or is willing to create one
- Has, or is willing to create, a Google or Apple ID (required to download the SMART-Health app from the Google Play or Apple App store)

4.3 Recruitment

SMART-Health will recruit a large number of adult participants of all ages, with and without a wide range of health conditions.

Participants may be recruited from across the UK but we will initially focus recruitment in South Yorkshire.

We will recruit from both NHS (participating NHS sites including secondary care hospitals (initially Sheffield Teaching Hospitals NHS Foundation Trust), primary care providers) and non-NHS settings (University staff, community groups, via direct approach to the population, etc).

Our recruitment strategies will include the following methods:

1) using social media, posters/flyers, advertising, presentations, and email lists of our institutions and partners (including but not limited to NHS organisations, community groups, Universities, Schools, Colleges) to directly invite potential participants to consider taking part.
2) via the direct care team of NHS patients (including but not limited to GPs, secondary care clinical teams, health and social care staff) will signpost potential participants to study information and the participant information sheet that invites them to consider taking part.
3) recontacting participants in existing cohorts, registries, studies or trials (if allowed by these studies' approvals) by email, post, or other method to invite them to consider participating.
4) using services such as the NIHR "Be Part of Research" [7] and SHARE [8] to contact people who have already expressed interest in taking part in research projects.

Participant information and consent to participate will be obtained using online forms (as used in Our Future Health) without the need for face-to-face meeting with a researcher. This will reduce barriers to participation.

However, some participants may choose or be asked to attend a face-to-face or online/telephone encounter, particularly for participants recruited during clinical care or where wearable devices are required to be provided to the participant. These interactions may take place at NHS settings such as NIHR Clinical Research Facilities or in community spaces.

Recruitment materials will signpost potential participants to a website that will be a single point of contact for all SMART-Health participants or potential participants.

The website will present the Participant Information Sheet (and allow this to be downloaded and printed). Potential participants can be sent a printed PIS if preferred. Potential participants will read this PIS which will include contact details for the SMART-Health team should they have any questions about the research.

A video with a transcript will also present the participant information sheet. This will be available via Youtube, our website and on the online recruitment portal.

The process of recruitment is shown below:

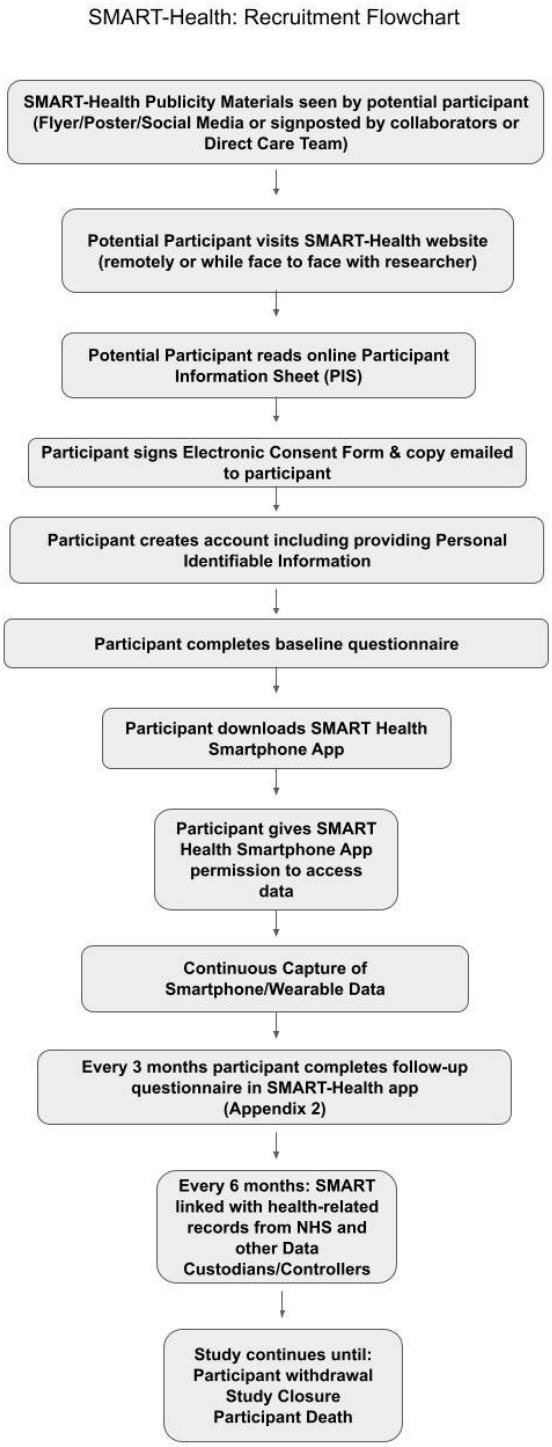


Figure: SMART-Health recruitment flowchart

4.4 Consent

Once the potential participant has read the PIS they will read and complete the consent form, either remotely or in the presence of a researcher or member of the direct care team.

The consent form will be completed using an online form on an appropriate secure platform such as RedCap. An electronic signature will be obtained (the participant will type their name and date).

If a participant is provided with a smartphone or wearable (such as a Fitbit or Apple Watch) by the SMART-Health study, participants will be required to complete the same user registration and accept the manufacturer's terms and conditions as for any user of that device. This will occur after consent is obtained, and the consent form will make this requirement clear.

Participants will receive a copy of their electronically signed consent form by email, or by post if they prefer.

5 Data to be obtained from participants

5.1 Baseline data to be obtained after consent

Once the participant completes the consent form, they will be directed to a secure online set of questions and questionnaires.

These are intended for completion at baseline, but we will recontact participants occasionally by email/text (no more than every 6 months) to ask if they need to update any that may have changed (such as address).

This baseline data includes

- Personal Identifiable Information (PII)
- Baseline Health Data
- Demographic Data

5.1.1 Personal identifiable Information (PII)

PII is required to create an account for each individual participant and to link their smartphone and wearable data with their health-related datasets.

PII is also required to recontact participants (to provide information to the participant and to approach to consider participation in other studies) and to allow scrutiny by research regulators and the study sponsor.

After providing consent, participants will be directed to a further secure online form (also hosted on a secure platform). Here a **pseudo-identification alphanumeric code** will be randomly generated and participants will be asked to provide Personal Identifiable Information including:

- Full name - Participant Forename, Participant Surname
- Gender (What sex were you registered with at birth? Female/Male/Intersex/Prefer not to answer)
- Date of birth
- Home address
- Email
- Mobile phone number
- NHS number if known (can be found in NHS app, hospital or GP letters, prescriptions or on medication boxes)

This and the pseudo-identification code will form the central recruitment log which will be stored with the highest security separately from all other data.

This recruitment log will have different access procedures to all other data and will be only accessible to restricted members of the SMART-Health team (expected to be a data manager and the PI, with access provided to the sponsor or appropriate regulatory authority on request).

PII will not be used in research however participant's age and socio-economic and environmental indices will be derived. De-identified derived data will be transferred to the separate database holding demographic data alongside the pseudo-identification alphanumeric code.

5.1.2 Baseline Health and Demographic Data

At baseline, a minimum dataset of baseline health and demographic data is required to allow some analysis of smartphone and wearable data prior to linkage with NHS data.

This data will be captured through a further secure online form and labelled with the participant's pseudo-identifier code. This will not contain any PII but is considered sensitive data and so will be held securely and separately to previously collected identifiable information. This pseudonymised data will be used in research and will be accessible to researchers linked with smartphone and wearable data and healthcare data.

Participants will be asked to answer all questions as soon as possible after providing consent. Reminders will be sent if answers are not completed (email/text). Where such prompts are used, the participant will be offered the opportunity to choose not to receive these so that they are not asked repeatedly for information they are unwilling to provide.

A key reason to seek medical information at this stage is to enable identification of people with specific health conditions or diseases within the overall cohort. It will be easier to identify them at an early stage rather than retrospectively using the linked NHS data.

Self-reported demographic data is required to characterise the diversity, representativeness and heterogeneity of the SMART-health cohort and its data. Demographic data is also necessary to understand the influence of these factors on patterns of smartphone and wearable data and to adjust for potential confounding factors when examining patterns of health.

Appendix 1 contains the text of the specific questions. The design of the web-based interface will follow guidelines to be maximally accessible [9].

The headings for the baseline survey questions (Appendix 1) are:

- **Questions about you and your household**
- **Questions about your health and wellbeing**

5.2 Smartphone and Wearable data

The smartphone and wearable data to be collected and the technologies used to collect this are detailed below.

Participant consent allows us to access data generated before the date of consent. On average this allows access to 2 years of retrospective data (usually the point of beginning to use the device).

5.2.1 Smartphone data

In the lifetime of SMART-Health, various technologies are likely to be used to acquire smartphone/wearable data. These will be subject to protocol amendments as needed. All technologies will undergo Information Risk Assessment by Data Protection Officers at the University of Sheffield if required.

All smartphone and wearable data will be securely transferred to the University of Sheffield Secure Data Services or other appropriate secure data environment. Prior to entering or once within the Secure Data Services, de-identification of smartphone and wearable data will be confirmed so that researchers who access this data cannot re-identify participants.

The initial SMART-Health smartphone app will be custom-designed for SMART-Health and used to gather participants' smartphone data.

Following obtaining consent and baseline questionnaire completion the participant will receive further instructions via email and directed to download the SMART-Health app from the Google Play (Android) or Apple App store (iPhone).

The app will collect data generated by users (questionnaires), and from connected wearables and the phone's internal sensors as follows:

- **Estimated step count** (derived from the phone accelerometer) at the highest sampling rate derived by the device.
- **Completion of follow-up survey questions** on such as life satisfaction, health status, physical activity, EQ-5D-5L, no more than every three months (see appendix 2 for details of follow-up questions)
- **OPTIONAL: GPS-derived location (continuous).** If permission provided by the participant, geolocation data will be mapped against existing public databases to derive non-identifiable features including: mobility (such as distance travelled per day), time spent in different types of transport, exposure to air pollution, traffic, green space, weather patterns, and other environmental health influences. **Participants must opt-in to collect GPS data and can participate without providing GPS data.** This is clearly explained in the PIS and consent form.

The user will be able to view on the SMART-Health app what data is being shared and summaries of their smartphone-derived data.

Data is transmitted from the smartphone to a UK-based secure cloud database using a unique token identifier. Data may be first transferred to secure servers in the UK before being securely transferred to the UoS Secure Data Service. For GPS data specifically, following feature extraction and transfer to secure data storage, raw geolocation data is deleted as per data processes defined below. Because raw GPS data is inherently at risk of re-identification, this data will never be made available to researchers and will only be accessible to a small number of the core SMART health team.

5.2.2 Wearable Data

Participants can join SMART-Health without owning or using a wearable device. In most cases we will then only collect the smartphone data above. However, for participants who use or are willing to use a wearable device connected to their smartphone, we will collect data from the wearable via the SMART-health smartphone app or via direct transfer from manufacturers' servers via the relevant

API (application programming interface). This will allow us to collect data from a wide range of connected wearables, including Apple Watch, Fitbit, or other wearable devices.

We will obtain the following wearable-derived data:

- **Estimated step count** at the highest sampling rate derived by the device (both phone and wearable will be providing this data although the two sampling methods will generate different data)
- **Sleep** (time of onset, offset, duration and where available sleep phases)
- **Heart rate** (at the highest sampling rate possible from the device)
- **GPS-derived location** (activated when the user triggers recording of an activity). Because raw GPS data is inherently at risk of re-identification this data will never be made available to researchers and will only be accessible to a small number of the core SMART health team. This team will use the raw GPS data to derive non-identifiable data as described in the section on smartphone data. **Participants will need to opt-in to collect GPS data.** They can participate in SMART-Health without providing GPS data, as explained in the PIS, and consent form.

All data streams will be time-stamped and labelled with non-identifiable metadata (what device was used, etc).

The user will be able to view on the app what devices are connected, what wearable data is being shared and summaries of their own data.

Data is transmitted from the smartphone to a UK-based secure cloud database using a unique token identifier. Data may be first transferred to secure servers in the UK before being securely transferred to the SMART-Health Secure Data Environment as described below.

5.2.3 Provision of smartphone and wearable devices to participants

Although we will initially aim to recruit people who already own smartphones and wearable devices, to improve accessibility and reduce digital exclusion we will seek funding to provide devices to some participants.

We have an initial supply of 250 Fitbit devices to provide to participants who wish to provide wearable data but do not have such a device.

We will apply for further funding to provide smartphones and wearables from other manufacturers so that a wider range of people can participate and to make our database as heterogenous and representative as possible.

To use a device provided by SMART-Health, the participant will usually need to create a user account and accept the usual manufacturers' terms and conditions for such devices. This is specified in the PIS/ consent form.

6 Linkage of Smartphone and Wearable data with healthcare records

NHS healthcare data will be requested from local NHS record systems, nationally held data and data custodians such as the NHS, UK Statistical Authorities and other UK bodies that store health information such as the Office for National Statistics, disease registries and GP records. Initially we expect to work predominantly with data from our partner the South Yorkshire Integrated Care Board.

Although, in general, only parts of individuals' health relevant records will be requested (particularly structured data such as diagnostic codes), consent will cover access to the full records. This includes past records, since these will help to characterise participants and to understand later health events more completely. This will relate to data captured in different settings: GP, hospital including A&E, dental as well as disease registries and occupational health records. The full records may also be required when necessary to verify the accuracy of data.

7 End of data collection

An individual participant's participation ends at the point of their withdrawal or death. The end of study is defined as the completion of any follow-up monitoring, data collection and analysis. We will reapply for ethical approval for the database every 5 years.

7.1 Withdrawal from study

Annual email updates will include a reminder that participants are able to withdraw will be sent to all participants.

Participants are free to withdraw at any time without giving reasons and without prejudicing any medical treatment.

Participants can withdraw by contacting us via our email, website or telephone.

Participants can also remain in the study but stop providing smartphone or wearable data by deleting the SMART-Health app. This will be made clear in the app and other materials.

Withdrawn participants will be provided with a contact point for further information about the study. Participants will be asked if they would like to provide the reason(s) why they have withdrawn consent, although this is voluntary. The record of withdrawal will be kept in a Withdrawal Tracking Sheet.

8 Safety

8.1 Assessment and management of risk to participants

This is a non-interventional observational study. The risk of serious direct adverse events is very low. Most participants will use personally-owned devices. Any devices provided to participants will be CE/CA marked consumer devices already on sale direct to consumers in the UK.

The most serious risk to participants is a data breach. The safeguards against such a breach and the steps we take to minimise the consequences of any breach are detailed in Section 9.

If any data breach or unexpected safety event occurs these will be notified to the study sponsor and relevant bodies immediately.

8.2 Incidental clinical findings

We have referred to the UKRI Framework on the feedback of health-related findings in research in considering and devising our policy.

We will not feed back health-related findings to the participants or their primary care physician because at present **there is no proven clinical utility** in such data. Smartphones and wearables are not medically approved devices and data from these devices cannot reliably indicate clinical findings that should be acted upon.

For example, heart rate accuracy from these devices is often poor, and heart rate will fall to zero when the device is not worn. It is not practical or desirable to act on such data by contacting the participant or their GP. Although some wearable device alert for signs of clinical issues (such as falls or irregular heart rhythms) we are not obtaining these alert data from the devices as we are focussing on the types of data detailed in section 5.2.

The participant information sheet and consent form make clear that the data is not being monitored and cannot be used in a way that replaces medical care. Participants will be reminded of the need to contact their usual health care providers as usual.

In addition, all healthcare data will already have been available to the participants' health care providers for an extended period of time prior to linkage meaning that no unknown incidental findings will be discovered that were not known to the direct care team.

9 Data Management and Access for research

9.1 Data flows including linkage to NHS data

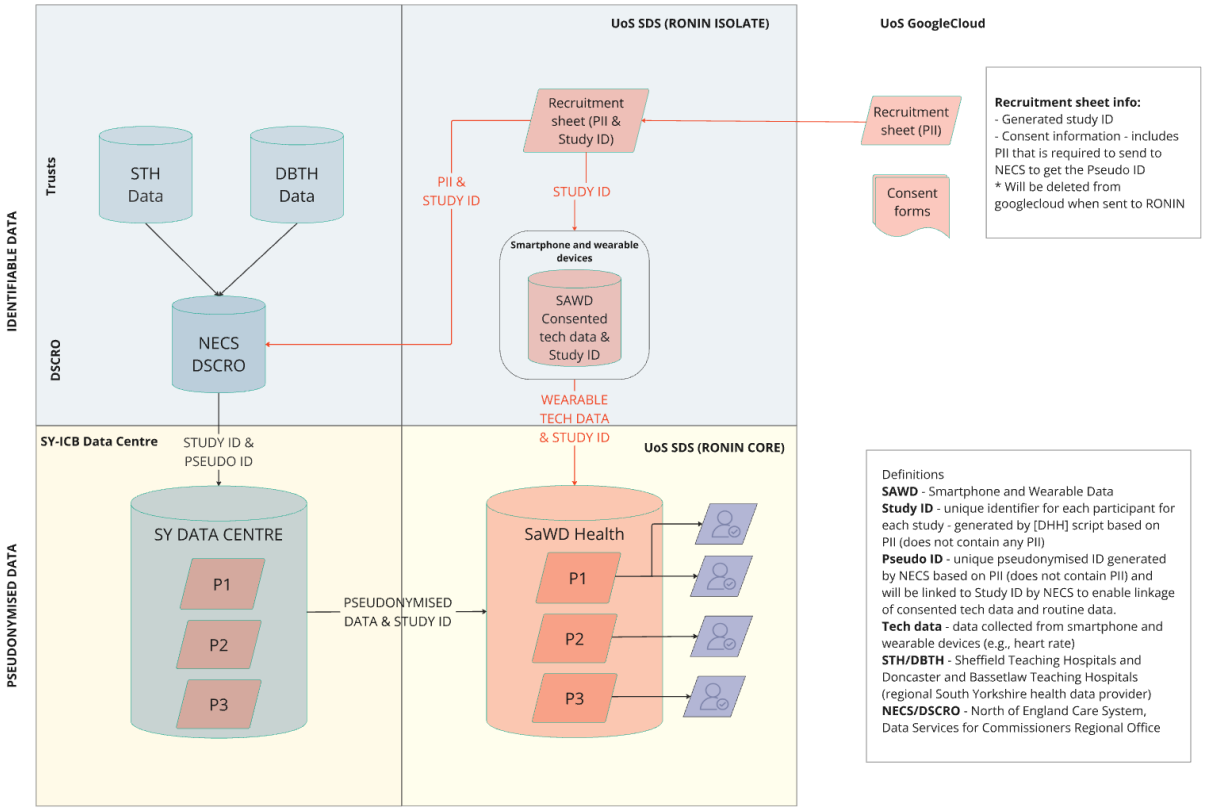
The SMART-Health Database will be hosted within an appropriately accredited Secure Data Environment, such as the University of Sheffield Secure Data Service.

This Secure Data Environment will ingest smartphone and wearable data and healthcare data, allow linkage of these and facilitate analysis of data "slices" (specific extracts of data and participants with de-identified characteristics).

This environment will have the highest appropriate levels of safeguards, transparency and access requirements to protect the security of the data while allowing the research community to access the data safely and perform analytics within the environment so that participants are not identifiable from research outputs. Data management will adhere to guiding principles that research data are findable, accessible, interoperable, and reusable (FAIR), as well as being attributable, legible, contemporaneous, original, and accurate (ALCOA). All research will be carried out in compliance with appropriate laws, rules, regulations, and guidelines applicable to the collection, use, handling, disposal, and processing of personal data. In particular, research will adhere to the provisions set out in.

- General Data Protection Regulation (GDPR)
- Directive 2006/24/EC of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communication services or of public communications networks

The proposed data flow model is illustrated in the Figure below.



A SMART-Health participant's completed consent form and the baseline health/ demographic data questionnaire will be captured within a UoS SDS hosted RedCap instance alongside a pseudo-identifier code. This log of the personal identifiable information (PII), and the self-reported research data will be securely stored separately within a secure environment such as the University of Sheffield SDS Platform. Access will be strictly limited to the smallest necessary members of the SMART-Health team. A data breach would reveal participant's identifiable details but not provide access to their smartphone, wearable or NHS data. Participants' age and socio-economic and environmental indices will be derived from their date of birth and postcode respectively and combined with the other baseline health and demographics data. Once de-identified, it will be transferred into a separate environment (UoS SDS Ronin Core) in preparation for linking with other research data. External researchers will never have access to PII.

The smartphone and wearable data captured by third party providers is described in Section 5.2. These data undergo initial automated preprocessing by those providers to generate the research data. All third party data processors will be assessed by the Information security team and approved by the Data Protection Officer at University of Sheffield. Smartphone and wearable data will also be ingested to UoS SDS Ronin Isolate from third party providers to the secure data environment accompanied with the pseudo-identifier code. Once the smartphone and wearable data is de-identified, it will be transferred into a separate environment (such as Ronin Core) in preparation for linking other research data.

- Raw or least-processed form GPS data since this is inherently identifiable. It will be securely stored and managed and non-identifiable measures will be derived (derived metadata [i.e. weather], type of location [home / work] or distance travelled per day) after which it will be deleted. A data breach of this database may provide access to a small amount of raw GPS

data (the data which has not yet been used to calculate de-identifiable measures or has not yet been deleted).
- Meanwhile, some smartphone/wearable device data contains metadata (such as the user-specified name of the device that generated the data). A data breach may expose a small amount of smartphone and wearable data that has not yet been de-identified by removal of metadata.

The **routinely-collected health data** of interest is described in Section 9.2. To allow linkage with research data, participant pseudo-IDs with a minimum of personally identifiable information will be shared with a trusted third party to allow linkage to be performed.

- For data obtained via the ICB the trusted third party will be the regional Data Services for Commissioners Regional Offices (DSCROs).
- For nationally held datasets an appropriate trusted third party will be selected such as Digital Health and Care Wales (which is able to link to all healthcare datasets, not just Welsh data).

Within the Trusted Third Party, NHS employed staff external to SMART-Health will extract NHS record data for SMART-Health participants and return the pseudonymised health record data with the pseudo-ID to the secure data environment (Ronin-core).

Upon receipt of a successful data access request and once the de-identified datasets are prepared within Ronin-Core, the SMART-health data management team will prepare extracts of the linked datasets to be made accessible to researchers within the Ronin-Core environment. More details of the data access process for researchers is detailed in Section 9. For these data “slices” a new pseudo-identifier code will be generated to protect the integrity of the rest of the SMART-Health database. No individual level linked datasets will leave the environment; only anonymous results of analyses of aggregated data will be exported after output checking.

9.2 Selection of NHS data linkages for all participants

The specific NHS datasets required for analysis will vary by the research question. We therefore seek consent and ethical approval that covers access to participants’ full past and future healthcare records. Data Sharing Agreements will be put in place with any appropriate healthcare and health-related Data Controllers/Custodians (such as local hospitals, primary care, NHS England, Subnational Data Environments, National Registries, etc).

Given that the initial recruitment will target residents in South Yorkshire and as a separate Data Sharing Agreement is required specifying each healthcare dataset to be obtained from each Data Custodian, we will initially request a minimum dataset for all participants from the South Yorkshire Integrated Care Board under a single initial Data Sharing Agreement through Connected South Yorkshire (ConSY).

The data to be initially requested on all participants will include previous and future:

- Ambulance Data
- Community Services Dataset (CSDS)
- Medicines Dispensed in Primary Care (NHSBSA Data)
- NHS Pathways Data Set
- Secondary Uses Service Dataset (SUS) - Hospital episode statistics data (HES)
- Primary Care
- Lab test results

- Diagnostic imaging dataset/Radiology Reports
- Maternity Services Data Set
- Mental Health Services Data Set
- Civil Registration of Deaths

Although all this data will be obtained and linked within the Secure Data Environment, researchers will only be provided access to the datasets pre-specified within the data access request form (including protocol and analysis plan) as approved by the Data Access Committee (Section 9.8).

For linkage to additional NHS datasets, the SMART-Health Steering Group will liaise with NHS and other data custodians to put in place appropriate data sharing agreements provided this is covered by participant consent.

NHS data will generally be requested on a six-monthly basis for all new participants and to update existing individual's dataset.

9.3 Security and audit measures to secure access to identifiable data

Identifiable data will be stored securely in distinct project spaces within a cloud-based secure data environment such as that managed by the University of Sheffield's Secure Data Service (SDS). This environment uses accredited Amazon Web Services (AWS) secure cloud storage and complies with NHS Digital's Data Security and Protection Toolkit standards. IT Services administrators have implemented robust policies to address network security (including external threats) and ensure timely software maintenance. Access to the environment is strictly limited to authorised users, with all accounts requiring multi factor authentication (MFA). MFA requires a username, password, and an external authenticator device for added security.

Identifiable or potentially identifiable data in SMART-Health is in three categories: personal information, healthcare data, and smartphone/wearable data. Each type is handled with specific security measures to ensure privacy and compliance.

- **Personal Information** such as name and NHS numbers, are stored separately in an encrypted database (e.g., secure Database in virtual machines, or REDCap) and do not leave the secure environment. During pseudonymisation, the SMART-Health team generates unique non-sequential Participant pseudo-IDs for individuals.
- **Healthcare Data** extracts, which may contain identifiers or identifiable information are securely stored and accessed only by the SMART-Health data management team and curated and cleaned of any identifiable information prior to any access by researchers. Researchers requiring such data must provide a detailed research purpose, which the Data Access Committee will approve (or reject) before granting access (see below)
- **Other Personal Information:** Postcodes and addresses are converted into pseudonymised Unique Property Reference Numbers (UPRNs) using geographical reference data from the Ordnance Survey. This conversion enables the grouping of individuals by household and access to information about residence types without exposing sensitive information. Using UPRNs ensures consistent referencing, simplifies location-based data analysis, and enables accurate data matching and sharing across systems, regardless of property management..

- **GPS Data** is accessible only to a limited number of approved SMART team members who are directly involved in the de-identification process. No external individuals or researchers are granted access to raw data.

Projects hosted within the SDS are regularly backed up in accordance with UoS SDS policies. To ensure consistency and transparency, data files and code are named systematically, and their storage locations are documented in a shared asset register, allowing all team members to locate datasets and relevant code easily. Access keys and passwords for storage locations within the SDS are securely stored in a KeyPass database. Where necessary, this information is shared only among authorised members of the SMART project team.

These measures, including data partitioning, pseudonymisation, encryption, controlled access, and regular backups, ensure robust security, uphold participant privacy, and effectively manage identifiable data within the SMART health database.

Further, SMART-Health will adhere to the guidelines set out in the central Information Security, Cyber Security and Data Protection policies outlined by the University. SDS has regular audits for their secure- cloud-based platform, both in line with the Audit requirements as complying with their ISO 27001 (2022) certification; as well as in line with the requirement of the Sensitive Research Data Services Information Management Committee.

SDS cloud platform has Data Backup, Encryption and Patching Policies, on top of central UoS Information Security policies. The central University of Sheffield Information Security Policy makes reference to handling information in line with the University of Sheffield Classification & Handling Scheme; data on the Secure Data Services platforms will be held in accordance with our own Data Classification Documentation.

9.4 Monitoring of the Database's systems and procedures

SMART Information Governance Team will conduct routine monitoring and audit of the database to ensure compliance with both study-specific and university policies. It will ensure the database is secured and analyse any potential data breaches.

On receipt of healthcare data, the SMART data management team will run validation checks to ensure the data received matches what was requested from the provider and aligns with data specifications for each dataset extract (e.g. relevant data dictionaries and technical output specifications for each dataset). The SMART data management team will then carry out data cleaning and standardisation work to bring all the datasets into a consistent format for inclusion in the SMART-Health research database.

The SMART-Health study steering group and University of Sheffield Secure Data Services will monitor all systems and procedures with an annual audit and review of SOPs with amendment as required.

9.4.1 Data handling and record keeping

Accumulating data will undergo central checks and monthly data management reports will be generated for resolution of any issues. Transfer of smartphone and wearable data into the secure data environment will be monitored for interruptions or technical issues.

Once data enters the Secure Data Environment it will be curated, metadata will be annotated, and de-identified by the SMART data management team.

Quality control procedures will be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly.

At the end of the life of the database, when all data has been coded, validated, and locked, a clean file will be declared and the final database will be locked.

9.4.2 Quality Control, Monitoring, Audit & Inspection

At regular intervals, analysis of data within the SMART-Health database will monitor recruitment rates and sources and provide an overview of active participants.

A data management team will undertake continuous data monitoring to ensure completeness and accuracy of clinical and digital data. The flow of screening, recruitment and follow-up of cohorts will be monitored regularly, such as monthly.

9.5 Access to Data for governance

Direct access to data (including Personal Identifiable Data) will be granted to authorised representatives from the Sponsor, funder, host institution and the regulatory authorities to permit study-related monitoring, audits and inspections.

9.6 Access to Data for research

At a point, determined by the study steering committee when sufficient participants have been recruited and data collected and linked, researchers will be invited to apply to access SMART-Health data extracts.

A SMART-Health Data Access Committee chaired by the CI or appropriate deputy will be responsible for considering, approving, rejecting, deferring and managing applications. This committee will comprise at least five members, and at least one lay member.

The process will follow the “Five Safes” framework [10] (Safe Data, Safe Projects, Safe People, Safe Settings, Safe Outputs).

Applicants (including University of Sheffield staff not in the SMART-Health team or representing the sponsor) will not have access to or process PII data.

9.6.1 Eligibility to apply to use SMART-Health

In order to be eligible to apply to access and use the data, an applicant must be employed by an organisation with a bona fide reason to conduct such research in the public interest. This is expected to usually be a Higher Education Institution, research Institute or healthcare organisation.

Where the primary applicant is not employed by University of Sheffield, a visiting academic agreement would be put in place and the named team on the project application should include a co-investigator employed at the University of Sheffield. This is in line with UoS Secure Data Services (SDS) requirements and ensures adherence to data security policies.

Representatives from voluntary, community, and social enterprises will be eligible to propose a research application in collaboration with a lead academic from University of Sheffield or Sheffield Hallam University.

An application including representation from a commercial partner will be eligible if the applicant team includes an academic employed by University of Sheffield or Sheffield Hallam University and if the University of Sheffield's due diligence considers that this does not pose a risk.

Applicants from academic, community and NHS organisations may be asked to provide costs on a cost-recovery model. Applicants from teams with commercial partners will be expected to provide costs at an enhanced rate, to be determined on a project-by-project basis and University of Sheffield commercial costing models.

All applicant teams will be encouraged to include public co-applicants or appropriate public/patient involvement and engagement.

9.6.2 Application process to use SMART-Health

Eligible researchers will apply to use SMART-Health via a secure online Data Access Request Form (DARF). This will include:

- Project Title
- Project Plain English Summary
- Confirmation of eligibility to apply (see above)
- Name of Principal Investigator of the Project
- Names, qualifications and CVs of applicant team members
- Records of required mandatory training (e.g. Good Clinical Practice, Information Security)
- Primary & secondary research question/s
- Details of alignment with SMART research priorities.
- Specification of smartphone and/or wearable datasets required
- Specification of linked NHS datasets required
- Research Methodology (Including usage of Artificial Intelligence, Machine Learning models)
- Data management plan of required datasets
    - Details on data processing required (including derivation of new variables)
    - Handling of missing data
- Statistical analysis (including prioritisation of outcomes).
- Letter of permission from local NHS R&D office to apply to use SMART-Health (for NHS researchers)

9.6.3 Review Process

Submitted applications providing the information above will be initially reviewed by at least two members of the SMART-Health Data access committee. If further information is required this will be sought from the applicant.

The review process will assess all applications against the following criteria:

- The project must address at least one of the 4 research questions within scope (section 9.6.5 below)
- The conduct of the research is in the public good.
- The data requested is available and appropriate to the research question.
- The requested data will also take into account the parameters outlined in the data sharing agreements with the various data custodians and organisations from which the health data has been obtained.

- The methodology including statistical analysis and statistical power is well-described and appropriate.
- The project should obtain an ethics self-declaration from the University ethics committee of the lead applicant's institution, based on their proposed methodology.
- Risks of bias and worsening of health inequalities have been considered and addressed.
- Due diligence on non-academic applicant organisations does not reveal reputational or other risk.
- There is sufficient capacity within the applicant team and SMART-Health team to perform data extracts and support the project.

At the committee meeting each application's two reviewers will present their recommendation to the entire committee who will also consider the application against the above criteria. The committee will be expected to reach a unanimous decision whether to approve, reject or defer. In the case of inability to reach a unanimous decision, at least a 70% vote in favour will be required to approve access. Researchers may apply to access only smartphone and wearable data without NHS record data, this will be reviewed by the Data Access Committee as for other applications.

Rejected or deferred applications may be allowed to resubmit a revised application that addresses the Data Access Committee's concerns.

For approved projects a Data Access Agreement will be put in place with the requesting organisation, and data access procedures followed to enable research to take place.

9.6.4 Record of projects using SMART-Health

A public record of all projects applying to and approved to use the database will be maintained including:

- Title
- a brief summary of its purpose
- Study Aims and importance
- dataset/s used, (including any sensitive data)
- name of the Chief Investigator
- the sponsor
- the location of the research
- the date on which the project was approved by the Research Database team
- Methodology
- Summary of data extract
- Implications and Impact
- any relevant reference numbers

9.6.5 Research questions, methodology and outcomes

We will allow access to the database for research that examines how smartphone and/or wearable data could be used to predict, prevent, diagnose or monitor human diseases.

Although the specific methodology will be dependent on each project, we seek generic ethics permissions to support the following research questions:

1) Identification of approaches to generate **meaningful, potentially clinically-informative parameters from raw or least-processed smartphone and wearable data**. This requires development and application of curation and analysis pipelines. This research is intended to discover generalizable, interoperable measures of health that could be tested for their ability to predict, prevent, diagnose or monitor disease
2) Correlation of parameters from smartphone and wearable data (including those developed in 1) with **present and/or past health status**. Comparison between groups of participants with specific current diseases and appropriate control groups will allow identification of different patterns of smartphone and wearable data in people with and without specific health conditions. This research is intended to **discover diagnostic signatures of smartphone/wearable data.**
3) Comparison of parameters from smartphone and wearable data (including those developed in 1) between groups of participants who later develop specific current diseases and those who do not develop these diseases. This research is intended to **discover predictive signatures of smartphone/wearable data.**
4) Correlation of parameters from smartphone and wearable data from people with specific health conditions with **adverse health outcomes and complications** (such as hospital admission, repeat operations or other complications) This research is intended to **discover strategies to monitor people with specific diseases for signs of response to treatment, adverse outcomes or complications.**

9.7 Archiving

Recruitment sites will be responsible for archiving any paper-based study documents after a minimum of 5 years from the end of the study. Destruction of documentation should be notified to the Sponsor.

Source data will be maintained at the local site of capture in a de-identified manner for a period of time stipulated by the local ethics committee (normally 5 years from the end of the study). Once this period of time has elapsed, the original de-identified dataset, and the study key code, will be destroyed.

Smartphone and wearable data captured within this database will be stored for up to 5 years and individual's health record data will be retrieved from the NHS records until the end of the study, their death or withdrawal.

10 SMART-Health Management and Coordination

The research database will be coordinated and managed by the SMART study team. This team consists of the team members directly involved with collecting data and managing the storage of data. The Database Management team will follow standard operating procedures to maintain documentation and undertake regular auditing and monitoring to ensure compliance.

The DMT will monitor progress of ongoing sub-projects, through periodic reporting.

Group	Members	Responsibilities

Study Steering Group	Stakeholders, CI, scientific advisors, Protocol Co-Investigators, lay representatives	Oversight of database, review updates on SMART-Health progress, internal sub-study Registration monitor all systems and procedures with an annual audit and review of SOPs with amendment as required.
Study Management Group	CI, data managers, specific members of the SMART-Health team	Overall project oversight, liaise with sub-cohort leads, participant withdrawal
Data Management Team	CI, data managers, data specialists, data engineers, software engineers	Prepare data extracts, check SDS training etc. Liaise with data services, On receipt of healthcare data, the SMART-Health data management team will run validation checks to ensure the data received matches what was requested from the provider and aligns with data specifications for each dataset extract (e.g. relevant data dictionaries and technical output specifications for each dataset). The SMART-Health data management team will then carry out data cleaning and standardisation work to bring all the datasets into a consistent format for inclusion in the SMART-Health research database.
Information governance team	CI, Secure Data Services Staff, Data Protection Officer	SMART-Health Information Governance Team will conduct routine monitoring and audit of the database to ensure compliance with both study-specific and university policies. It will ensure the database is secured and analyse any potential data breaches.
Data Access committee	CI, Data managers, Secure Data Services Staff, Lay representatives, Co-investigators	review and approve data access requests

Table : SMART-Health management structures

11 Ethical and Regulatory Considerations

11.1 Investigators' Responsibilities

Investigators are responsible for performing the study in accordance with the UK policy framework for health and social care research the guidelines of the UK Health Research Authority, the Human Tissue Act, International Conference on Harmonisation Good Clinical Practice (ICH-GCP) guidelines, the Declaration of Helsinki guidelines (www.wma.net), and the General Data Protection Regulation (GDPR 2018). Investigators are required to ensure compliance to the protocol, Case Report Forms and Standard Operating Procedures. Investigators are required to allow access to study documentation or source data on request for monitoring visits and audits performed by any regulatory authorities.

The study will be conducted according to the declaration of Helsinki, Good Clinical Practice (GCP) standards and The European Code of Conduct for Research Integrity[24].

The Principal Investigator (or an appropriate member of the research team) is responsible for reporting protocol deviations/violations.

11.2 Study Sponsor

The University of Sheffield will be the sponsor and Data Custodian.

11.3 Ethics

The protocol will be approved by the University of Sheffield research governance team prior to commencement of recruitment.

Before the start of recruitment, a favourable opinion will be obtained from the Health Research Authority research ethics committee (REC) for the study protocol, participant information sheet, consent forms and other relevant documents. Substantial amendments that require review by local REC will not be implemented until that review has been completed and mechanisms are in place to implement at site. All correspondence with the local REC will be retained and sent to the sponsor.

Where recruitment occurs in clinical centres other than Sheffield Teaching Hospital, each centre will be responsible for the submission and approval of the study protocol to the relevant local ethical committees.

The Chief Investigator is responsible for producing the annual reports as required and to notify the REC of the end of the study. If the study is ended prematurely, the Chief Investigator will notify the REC, including the reasons for the premature termination. Within one year after the end of the study, the Chief Investigator will submit a final report with the results, including any publications, to the REC.

11.4 Study Funding

Funding for this research database was obtained from the EPSRC/UKRI Digital Health Hub pilot scheme, awarded in open competition to the University of Sheffield with Sheffield Hallam University as a partner.

11.5 Portfolio adoption

We will apply for SMART-Health to be adopted by the NIHR CRN portfolio as it fulfils the stated criteria [11]

Studies that have undergone ethical review as tissue banks or databases, but which meet the above definition of research and all other aspects of the Eligibility Criteria, can be supported by the RDN, if the following are satisfied: (1) The research questions and anticipated outcomes are clearly

stated; and (2) The research methodology to be used (in addition to methods of sample collection / processing / storage) are clearly described; and (3) The outcome(s) can reliably be extrapolated from the subjects who participated to a broader patient population and a broader range of clinical settings; and (4) Evidence is provided to confirm that funding secured covers all research costs as well as sample collection / processing / storage.

12 Team Expertise

Chief investigator: Professor Tim Chico is Professor of Cardiovascular Medicine at the School of Medicine and Population Health at the University of Sheffield and an honorary consultant cardiologist. He is Director of the South Yorkshire Digital Health Hub and Associate Director of the British Heart Foundation Data Science Centre, where he leads the Smartphone and Wearable data theme.

The SMART-Health research database team is comprised of internationally-leading academic and clinical researchers from the University of Sheffield, Sheffield Hallam University Sheffield Teaching Hospitals NHS Trust. We have extensive expertise in conducting routine health data research, deploying digital healthcare technologies, implementing state of the art data governance infrastructures and scientific computing.

The team includes members of the Advanced Wellbeing Research Centre (SHU), UKRI EPSRC South Yorkshire Digital Health Hub (UoS, SHU), Data Connect (UoS) and INSIGNEO Institute for insilico Medicine (UoS). We have links to the NIHR Yorkshire & Humber NIHR Applied Research Collaborative, HDRUK Northern Partnership projects, UoS Healthy Lifespan Institute, and the British Heart Foundation Data Science Centre. We have strong relationships with South Yorkshire Integrated Care System, Health innovation South Yorkshire and the South Yorkshire Mayoral Combined Authority.

13 Patient and Public Involvement and Engagement

Patient and Public Involvement and Engagement (PPIE) in SMART-Health is essential to facilitate recruitment and consent. Service design, conduct, governance, and research must be transparent and supported by patients and the public. Such support is also essential for diverse participant recruitment needed to generate heterogenous datasets

The SMART-Health protocol and participant facing materials have been reviewed by the Patient Public Panel of the British Heart Foundation Data Science Centre and amended in line with their advice and comments.

At least two lay members will sit on the SMART-Health steering committee to provide ongoing input.

The Data Access Committee will have at least one lay member.

We will also conduct regular engagement events with patients and the public and seek feedback from potential participants and participants.

We will keep an action log of changes made in response to feedback from patients and the public.

14 Indemnity

Indemnity to meet the potential legal liability of the sponsor for harm to participants arising from the management and conduct of the research will be provided by the University of Sheffield.

15 Dissemination

Dissemination of results derived from SMART-Health is crucially important to reach a long-lasting impact. SMART-Health has multiple measures in place to maximise dissemination of the results, emphasising a stakeholder-driven dissemination strategy and an Open Access policy. The study will be listed on the clinicaltrials.gov registry and included on the NIHR Clinical Research Network Portfolio. Relevant stakeholders for SMART-Health include the scientific community; patients and patient organisations; health care professionals and public health authorities; pharmaceutical and associated industries; regulatory bodies; and the general public. Scientific dissemination will take place through peer-reviewed publications in scientific journals and presentations at scientific conferences. The results of the study will also be disseminated through popular-science and professional publications in a variety of trade journals and magazines. The wider audience will also be kept appraised of the study results through the SMART-Health website, social media, newsletters, press releases and project videos. In addition, participants will receive feedback in the form of a newsletter.

All draft research outputs will be checked to ensure participants are not re-identifiable. Only the combined research outputs from many participants will be disseminated.

15.1 Authorship eligibility

The study is expected to result in a large number of primary papers and a wide variety of secondary papers, opinion pieces and other outputs. Emphasis will be placed on ensuring early career partners in particular benefit from their contribution. Authorship on publications and presentations will adhere to the ethical guidelines for authorship on scientific output as recommended by the ICMJE (International Committee of Medical Journal Editors).

References

1. Deloitte's digital Consumer Trends: UK insights. *Deloitte United Kingdom* https://www.deloitte.com/uk/en/Industries/tmt/research/digital-consumer-trends.html.
2. Allen, N. E. *et al.* Prospective study design and data analysis in UK Biobank. *Sci. Transl. Med.* **16**, eadf4428 (2024).
3. Why diversity is essential to our mission. *Our Future Health* https://ourfuturehealth.org.uk/our-research-mission/why-diversity-is-essential-to-mission/.
4. All of Us Research Program Investigators *et al.* The 'All of Us' Research Program. *N. Engl. J. Med.* **381**, 668–676 (2019).

5. Barker, J. *et al.* Physical activity of UK adults with chronic disease: cross sectional analysis of accelerometer measured physical activity in 96 706 UK Biobank participants. *Int. J. Epidemiol.* **48**, 1386 (2019).
6. Sudlow, C. Uniting the UK's health data: A huge opportunity for society. Preprint at https://doi.org/10.5281/ZENODO.13353747 (2024).
7. TrialBlazers. *Be Part of Research* https://bepartofresearch.nihr.ac.uk/.
8. SHARE. *Register for SHARE* https://www.registerforshare.org/.
9. Web Content Accessibility Guidelines (WCAG) 2.1. https://www.w3.org/TR/WCAG21/.
10. UK Data Service. What is the Five Safes framework? *UK Data Service* https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/ (2021).
11. Study eligibility for Clinical Research Network Support - FAQs. https://www.nihr.ac.uk/study-eligibility-research-delivery-network-support-faqs#:~:text=Studies%20that%20have%20undergone%20ethical,have%20not%20previously%20been%20stored).
12. Liu, J. J. *et al.* Digital phenotyping from wearables using AI characterizes psychiatric disorders and identifies genetic associations. *Cell* **188**, 515–529.e15 (2025).

Change log

Date	Change	Significance
22/5/24	Flowchart updated (from SaWD-Health to SMART-Health)	None