+
Our
Future
Health

Our Future Health
Cloud TRE
Procurement – Cross-Lot Requirements

Version: 03 March 2022

Contents

1 Programme summary 2
1.1 Lot summary 2
1.1.1 Lot 1 - Cloud 2
1.1.2 Lot 2 – Trusted Research Environment 2
1.1.3 Lot 3 – Researcher Billing 2
2 Contract period 2
3 Background 2
3.1 Key assumptions 3
3.2 Architecture overview 3
3.2.1 Hosting of components 4
4 Responsibilities 5
4.1 Integration responsibilities 5
4.2 Operational responsibilities 5
4.3 Reporting and data requirements 6
4.4 Security incidents 6
4.5 Researcher account responsibilities 6
5 Project scaling 7
5.1 Participants 7
5.2 Researchers 7
5.2.1 Researchers Scenario 7
5.3 Data storage 7
5.3.1 Genetic data 7
5.3.2 Linked data 8
5.3.3 Participant generated data 8
5.3.4 Size of data 8
5.3.5 Data scenarios 8
6 Security 9
6.1 Incident response and monitoring 10
6.2 Ongoing review 11
7 Project Management 11
7.1 Milestones 12
8 Appendix 1: Architecture Overview 14

+
Our
Future
Health

1 Programme summary

Our Future Health aims to be the UK's largest ever health research programme. It is designed to help people live healthier lives for longer through the discovery and testing of more effective approaches to prevention, earlier detection, and treatment of diseases.

Our ambition is that five million adults across the UK will volunteer to take part in Our Future Health. Participants will be asked to provide information about their health and lifestyles and a small blood sample. We will ask participants for permission to combine the information and samples that they give us with existing information about them, including their health records. Participants will also be given the option to take part in additional research studies.

Combining these multiple sources of health and health-relevant information, including genetic data, will create an incredibly detailed picture that truly represents the UK population. Researchers will use this information to make new discoveries about human health and diseases.

1.1 Lot summary

1.1.1 Lot 1 - Cloud

Our Future Health aims to procure a secure and scalable public cloud infrastructure to support a health research programme. The selected Supplier will be one of our most important technology partners as we build and scale our systems. The cloud infrastructure will underpin websites and apps used by our participants, a data store and data processing systems, a Trusted Research Environment (TRE) and associated systems used by researchers from around the world, and underlying systems to operate and monitor the programme.

1.1.2 Lot 2 – Trusted Research Environment

Our Future Health aims to procure a Trusted Research Environment (TRE) to enable registered researchers to access and use data within a single secure environment to conduct approved research studies. An important measure of the success of Our Future Health is the research that is conducted with the data and the cohort, together with our commitment to robust security, strict governance, and participant confidentiality.

1.1.3 Lot 3 – Researcher Billing

Our Future Health is planning to procure a Researcher Billing Service to allow researchers using our Trusted Research Environment (TRE) to pay for their usage of cloud services including compute and storage.

2 Contract period

The initial contract period for each Lot shall be three (3) years with available extension options of twelve (12) months plus twelve (12) months.

3 Background

During 2021, Our Future Health ran pilot studies to test the viability of various ways of recruiting participants and collecting biological samples. The pilot studies were deployed on a minimal set of technology platforms. The aim of the pilot studies was to try to understand how participant

recruitment would fundamentally work and then take the learnings into the main study. The pilots did not generate genetic data from biological samples but ended with the storage of biological samples. For the main study, participant samples will be collected and shipped to one of the four processing laboratories across the UK – England, Scotland, Northern Ireland and Wales. Samples will then be shipped to a single central laboratory where they will undergo genotyping.

In 2022 we will be moving to a more scalable set of platforms to allow the programme to scale up. Our ambition is to recruit 5 million people by 2025. In addition, during 2022 we will be launching our TRE that will allow researchers from around the world to work with the Our Future Health data in a highly secure environment.

3.1 Key assumptions

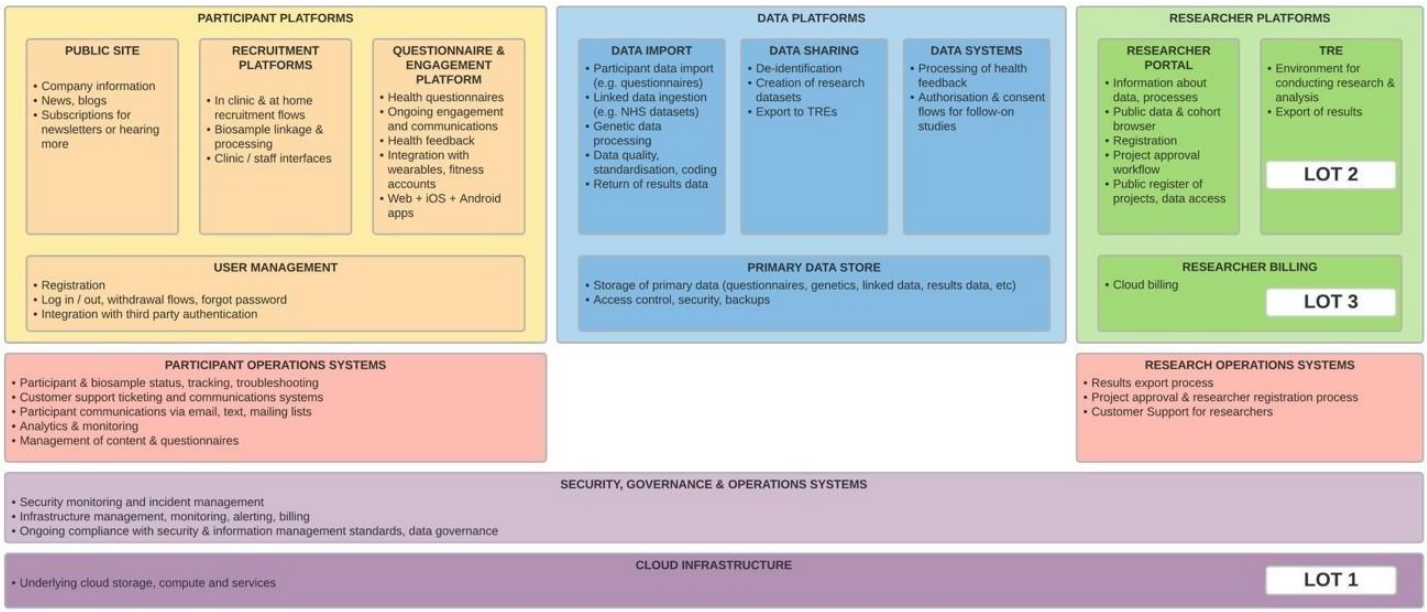
There are some underlying principles to guide this procurement:

- we will use public cloud
- our participant data will be hosted within the UK
- security of the cloud environment is a very high priority

Our rationale for UK hosting, particularly the primary data store that we explain below, is to minimise data governance complexity associated with international transfers.

3.2 Architecture overview

The diagram below shows the main components within our technology stack and their relationship to the lots in this procurement (See also Appendix 1 of this document for a larger version).



3.2.1 Hosting of components

This table shows how components are currently set up within our minimum viable product (MVP) pilot systems, and how we anticipate they will be provided as we move to higher scale platforms, including which components are expected to move into the new procured cloud.

Components	Current provision	Anticipated provision at scale
Public site	The pilot is using a platform-as-a-service (PaaS).	We will retain this arrangement as is.
Recruitment platforms	The MVP product is a bespoke mobile web app and back end	This will remain a bespoke app and will be hosted in the procured cloud.
Questionnaire & engagement platform	The MVP product is a bespoke mobile web app and back end	We may be procuring this as Software-as-a-Service (SaaS) with bespoke customisation. The hosting will be subject to negotiation with the selected partner.
Primary data store	Cloud databases	Storage systems in procured cloud.
Data import, data sharing, data processing	During the pilot studies, minimal processing takes place.	Bespoke software and tools running in the procured cloud.
Researcher portal	This does not exist yet.	This could be a platform we purchase, or it could be bespoke software hosted in the procured cloud.
TRE	The pilot is using an NHS environment, but this is only open to internal staff.	A TRE platform will be hosted in the procured cloud.
Participant operations systems	The pilot is using a variety of standard Software as a Service providers	We will retain CRM and Customer Support as is and review requirements during 2022.
Research operations systems	These do not exist yet.	This could be a platform we purchase, or it could be bespoke software hosted in the procured cloud.
Security, governance and operations systems	A variety of tools deployed in the cloud.	Tools running in the procured cloud.

4 Responsibilities

This document includes requirements that apply across the lots of this procurement and those relating to integration and cooperation between Suppliers once contracts are awarded.

The following abbreviations will be used:

- C (Cloud): Service Provider of Lot 1
- T (TRE): Service Provider of Lot 2
- B (Billing): Service Provider of Lot 3
- O (Our Future Health)

4.1 Integration responsibilities

Once contracts are awarded for all three lots, we will need to begin integration across the Suppliers, and this will be ongoing while the system is live and recruiting participants.

This table shows which Supplier will be accountable for each of the integration activities, and who will be responsible for any technical work. We assume that integration activities will require time and work from both Suppliers involved. The accountable Supplier should provide an indicative integration cost for integration with Suppliers of other lots, where applicable, and to Our Future Health, with reasonable assistance provided by their integration partner.

	Responsible	**Accountable**	**Consulted**
Integration of B with C	B, C	B	O, T
Integration of B with T	B, T	B	O, C
Deployment of T on C	T, C	T	O, B
Integration of T with O's systems and processes such as primary data store, access process, airlock process, researcher registration process	T, O, C	T	B

4.2 Operational responsibilities

The live services will have dependencies. T will run on C with connection to O for example. Here we list the technical support responsibilities for faults and incidents within the live services provided by each Supplier.

Live services provided by	**1st line support**	**2nd line**	**3rd line**
T	T If requests are escalated to T via O researcher helpdesk, then T will continue with direct end user contact	T	Ability to contact C, B, O support

B	B If requests are escalated via O researcher helpdesk to B, then B will continue with direct end user contact	B	Ability to contact C, T, O support
C	C	C	C, O
O	O	O	Ability to contact C, T, B support

Our Future Health will run a helpdesk for researchers to triage their requests. Requests that relate to technical issues will be raised with the appropriate process above, at which point the Supplier service will have direct contact with the researcher.

4.3 Reporting and data requirements

We will require regular reporting from the support function of each Supplier to include incidents raised, response times, fix times and other commonly used metrics.

Where a Supplier has direct contact with a researcher using the platforms as part of a support request, we require visibility of the communication and history of the support ticket.

4.4 Security incidents

Each Supplier must provide a capability to be contacted and contribute to an incident response team in the event of a security incident. Depending on risk and impact, the team will draw members from all Suppliers from this procurement, Our Future Health and Our Future Health's security supplier, and other platform or supplier representatives. Each Supplier must provide a capability to escalate incidents and trigger this process, to be available 24/7/365.

Each Supplier will provide their Incident SLAs with detailed information on:

- Incident tiering
- Incident response times
- Incident emergency contact information

4.5 Researcher account responsibilities

We assume T will be the owner of processes around account creation and deletion in the relevant systems, and delegate aspects to C and B as needed, since T will be the primary interface that a researcher uses.

The accountable Supplier must take delivery responsibility, work with the partners listed as responsible for any required integration.

	Responsible	**Accountable**	**Consulted**
Processes such as researcher registration and study approval	O	O	T, C, B

Creation of new account for researcher or entire new project and Workspace as needed across C, T, B systems	B, C, O	T	
Removal of individual researcher account or entire project and Workspace from C, T, B systems when required	B, C, O	T	
Tracking of researcher cloud costs and resulting billing activities	T, C	B	O

5 Project scaling

5.1 Participants

To guide planning assumptions around scale, Suppliers should plan for linear growth in participant numbers starting from May 2022 and reaching 5 million by November 2025.

Please note that actual growth may be different.

5.2 Researchers

A phased approach to researchers as users of the TRE is expected. We do not yet have estimates of how quickly we will be adding researchers to our platform. As a planning assumption, Suppliers should assume that we will have 200 researchers in 2023, 500 in 2024 and 1500 in 2025.

5.2.1 Researchers Scenario

The scenario we would like Suppliers to use as a planning assumption and for any pricing calculations related to Lots 2 and 3 is: 100 Workspaces, 100 projects, 100 research groups / institutions and 500 individual researchers. We do not have estimates for how much cloud compute or storage researchers will use.

Please note that actual growth may be different.

5.3 Data storage

5.3.1 Genetic data

The initial Our Future Health funding will enable the 5 million participants' samples to be processed via a custom single nucleotide polymorphisms (SNP) array. This genotype data will be imputed using a suitable reference panel, and both data sets will come back to our cloud environment for storage. This process will happen iteratively as the data are generated over time.

There is a possibility over time that additional funding may allow whole exome or genome sequencing of samples for all, or a subset of the participants.

5.3.2 Linked data

Our Future Health seeks to link to participants' clinical data from multiple sources, for example NHS primary or secondary care records.

The size and depth of linked clinical data is expected to grow significantly as recruitment scales and health records grow for each participant. In the future, and subject to additional consent from participants, we may add:

- medical imaging data
- further linkage of health-related data

5.3.3 Participant generated data

Participants will complete an onboarding questionnaire. In the future, and subject to additional consent from participants, we may add:

- activity tracking and/or wearable data
- other forms of participant generated data

5.3.4 Size of data

To aid with planning, we will use these orders of magnitude as estimates:

- NHS data – 20MB per participant
- genetic data (SNP) – 50MB per participant
- genetic data (whole Exome) – 7GB per participant
- genetic data (whole Genome) – 100-200GB per participant (model below assumes 100GB)

Note that we cannot guarantee that whole exome or genome data will ever be available; we are showing estimates here purely so that Suppliers can see a range of possibilities.

Participant health data will generally accumulate over time but rarely change.

5.3.5 Data scenarios

To help with understanding the necessary scale, we have created a few example scenarios. The scenarios are not intended to be an accurate projection of our data storage and workloads; however, we trust that they provide a valuable approximation.

Data is split into two high level types - 'relational' and 'object' - based on its predicted access patterns. 'Relational' includes clinical / questionnaire / participant data, 'object' includes genetic data / images / scans.

To help with planning assumptions we are using this model of how our service might scale in terms of the number of participants, although our actual growth may be different. All values in the tables below are given in terabytes (TB).

5.3.5.1 Base Scenario

Year	2022	2023	2024	2025
Relational data incoming (TB)	4.9	19	35	39

Relational data stored at year end (TB)	9.8	47	120	200
Object data incoming (TB)	15	56	130	190
Object data stored at year end (TB)	15	71	201	391

- Relational data requirements:
  - periodic bulk ETL-type processes, e.g., transformations of the data from one schema to another
  - not random access, i.e., we do not anticipate supporting participant website requests where any given record may be fetched at any moment
  - periodic exports to serialisation format
- Object data requirements:
  - periodic exports

5.3.5.2 Extended scenario 1 - storing whole genome

As previously mentioned, in the future it is possible we will store whole genome sequences for all participants, which means total data size increases by several orders of magnitude. In this scenario it is assumed the sequencing would start in 2023 for all participants.

Year	**2022**	**2023**	**2024**	**2025**
Object data incoming	0	120,000	180,000	200,000
Object data stored at year end	0	120,000	300,000	500,000

5.3.5.3 Likely scenario for initial beta test

When our TRE begins a limited beta testing period, we expect that we will only be releasing participant questionnaire data, with a size of this order:

Relational data stored	0.1
Object data stored	0

Subsequent data releases will follow, most likely with initial Hospital Episode Statistics linkages for England, and SNP data for a portion of the cohort.

6 Security

Suppliers are required who can work proactively with Our Future Health to support our compliance with strict data governance requirements for working with sensitive health data,

including GDPR, Cyber Essentials Plus, the NHS Data Security and Protection Toolkit, ISO27001 and ISO 27701.

Financial reviews for Suppliers shall be conducted by a qualified professional. The lack of a review must be justified for example where the supplier is clearly understood to be solvent. This review should include a) the value of the contract; b) the terms; c) the duration; d) Supplier's pricing history.

All Suppliers shall maintain formal ISO27001 certification through a recognised regional body, or demonstrate compliance with an equivalent standard or guidance where controls meet the intent of those in the ISO27001 Annex A control list.

All Suppliers shall attain certification with Cyber Essentials Plus, through an accredited certification scheme, within 6 months of contract award. The scope must include all assets comprising the systems delivering the requirements of this procurement, and any assets within the Supplier's own infrastructure having an effect on confidentiality, integrity or availability of those systems.

In addition, any Supplier's processing payment information must be PCI-DSS compliant and FCA registered as a minimum.

We prefer that Suppliers are also able to demonstrate accredited certification to the ISO9001 (Quality Management Standard).

Any deviation from this control requirement must be supported by a documented risk assessment and approval from the Our Future Health risk owner.

Data controller and data processor relationships require Our Future Health to ensure compliance with the GDPR article 28 where a processor will process data on our behalf. For the avoidance of any doubt storing of data is considered processing.

6.1 Incident response and monitoring

In the event of the invocation of a formal incident response:

- current communication plans for the Supplier shall be updated to include specific notifications to the Our Future Health Security Incident Response Team (SIRT), to an agreed SLA.
- Access and support will be provided to any engaged Digital Forensics and Incident Response (DFIR) functions (either provided by Our Future Health or an approved third party), in order to conduct root cause analysis and forensic investigation if needed.
- in the case of key staff who are allocated to work with Our Future Health, the Supplier's JML (Joiners, Movers, Leavers) process will include notification of key staff changes to Our Future Health.

As part of the Supplier's threat and vulnerability scanning, a process shall be put into place to escalate emerging threats to Our Future Health to plan appropriate responses if there are elements of shared responsibility.

6.2 Ongoing review

The Supplier will agree to monitor and develop the security program, based on changes and revisions to the applicable standards, regulation or contractual requirements. Where changes require new controls or alter the current controls implementation, the Supplier will engage with Our Future Health to update security assurance information.

A formal annual security review will be included as part of ongoing account management, to include:

- a review of prior incidents and their outcomes
- planned events or changes that may impact security of the services in the next 12 months

7 Project Management

Each Supplier will need to participate in a project management process to ensure a successful launch of these services. An Executive Steering Group involving members of the Our Future Health team (and others as Our Future Health may deem to invite) and appropriate members of the Supplier/s shall meet monthly during the initial period (from contract award to the end of the TRE Beta period) and quarterly thereafter to oversee satisfactory progress, timely implementation, operational performance, and future development plans (considering and prioritising feedback from the Our Future Health research community).

A regular cadence of meetings, at least weekly, is expected between a working group of relevant individuals from Our Future Health and the Suppliers from award of contract to beta launch (30th September 2022).

Each Supplier shall assign nominated project manager/s with relevant qualifications and experience of delivering similar projects, who will act as the principal point of contact for management of the implementation project including:

- delivery of all project phases and any subsequent releases
- ingestion of Our Future Health data,
- coordination of data releases
- project setup activities
- development of project documentation
- oversight of the platform implementation
- development of Standard Operating Procedures for service delivery.

The Suppliers shall have a defined approach to project management, either via the Supplier's own project management framework (evidenced with reference to their successful use on projects of a similar scale and complexity) or derived from an industry standard project management framework (e.g., Prince2).

Any dependencies on Our Future Health infrastructure, systems, or staff that need to be met to facilitate implementation and/or use of the service shall be identified.

7.1 Milestones

Timeliness of implementation is of critical importance to Our Future Health. The following timelines provide an overview of expected phases. Key milestones are commitments that Our Future Health has made and will also be referenced in supplier contracts.

Milestone	Deliverables	Description	Date	Key Milestone
1	Contract award		Effective date	No
2	Project kick off	Supplier responsible for Project Definition Workshop to inform completion of Detailed Implementation Plan	One week after Effective date	No
3	Delivery of implementation plan		Two weeks after Effective date	No
4	Alpha TRE hosted in Cloud	● TRE (Lot 2) running on cloud (Lot 1) ● On-boarding and access to the TRE for Customer staff ● Ingestion and availability of synthetic data for preliminary configuration.	As soon as possible after Effective Date, no later than 31 July 2022	No
5	Beta TRE hosted in Cloud	● Ingestion and availability of Customer data to TRE (questionnaire and demographic data only) ● Delivery of the functionality required as set out in the Specification ● Successful completion of testing as set out in Section 5.3 of the Specification and Schedule 5 (Testing Procedures) of the Agreement ● On-boarding and access to the TRE for additional approved researchers as identified by the Customer ● Researcher Billing MVP functional for approved researchers with required integration between all Lots	30th September 2022	Yes
6	Phase 1 TRE	● Accreditation of TRE ● Ingestion and availability of Our Future Health data to TRE (questionnaire and demographic	31st December 2022	Yes

		data, linked health data (if available) and initial SNP arrays) ● Triage of Beta bugs and backlog for launch. ● General availability and access for all approved registered researchers (with on- boarding determined by User demand). ● Full Researcher Billing functionality live		

8 Appendix 1: Architecture Overview

This is also provided separately as ‘Our Future Health - Cross-Lot Requirements (Appendix 1 Architecture Overview) - FINAL 03 March 2022’.