Data Protection Impact Assessment - NHS England OpenSAFELY COVID-19 Service

Document filename:	**Data Protection Impact Assessment – NHS England OpenSAFELY COVID-19 Service**	
Directorate / Programme	**Product delivery strategy and operations**	
Document Reference (previous reference number: IG20200468)		
Information Asset Owner		Version **1.0**
Author Redacted		

Document Management

Revision History

Version	Date	Summary of Changes
0.1	23/03/2023	Initial draft of document in new template
0.2	28/03/2023	Second draft for wider delivery group review
0.3	28/03/2023	Third draft for wider delivery group review
0.4	05/05/2023	Updated draft for Interim Data Advisory Group discussion
0.5	16/05/2023	Updated draft for OS/NHSE discussion
0.6	17/05/2023	Updated risk table and risk section
0.7	25/05/2023	Updated to reflect Michael Chapman's feedback and incorporate clarifications from Bennett Institute OpenSAFELY team
0.8	25/05/2023	Updated following final review with IG colleagues. Approved for submission to DPO for review
1.0	02/06/2023	DPO Approval received. Awaiting feedback on ICO Consultation approach and addition of signatures for DPO. IAO signatory also to be confirmed

Reviewers

This document must be reviewed by the following people:

Reviewer name	Title / Responsibility	Date	Version
Eva Simmonds	GP Data Programme Director	28/03/2023	V0.3
Louise Whitworth Woodhead	Associate Director of Information Governance	25/05/2023	V0.8
Richard Langley	Head of Information Governance	25/05/2023	V0.8

Approved by

This document must be approved by the following people:

Name	Title / Responsibility	Date	Version
Jon Moore	Data Protection Officer	02/06/2023	V1.0
Garry Coleman	SIRO	30/05/2023	V0.6
Phil Koczan	Deputy Caldicott Guardian	30/05/2023	V0.6
Michael Chapman	Director of Research and Clinical Trials	25/05/2023	V0.6

Document Control:

The controlled copy of this document is maintained in the NHS Digital corporate network. Any copies of this document held outside of that area, in whatever format (e.g., paper, email attachment), are considered to have passed out of control and should be checked for currency and validity.

1. Contents

2. Purpose of this document

A Data Protection Impact Assessment (DPIA) is a useful tool to help NHS England demonstrate how we comply with data protection law.

DPIAs are also a legal requirement where the processing of personal data is *"likely to result in a high risk to the rights and freedoms of individuals".* If you are unsure whether a DPIA is necessary, you should complete a DPIA screening questionnaire to assess whether the processing you are carrying out is regarded as high risk.

By completing a DPIA you can systematically analyse your processing to demonstrate how you will comply with data protection law and in doing so identify and minimise data protection risks.

3. Background

From 1 February 2023, NHS England has assumed responsibility for all activities previously undertaken by NHS Digital. This includes running the vital national IT systems which support health and adult social care, as well as the collection, analysis, publication, and dissemination of data generated by health and social care services. The statutory functions of NHS Digital transferred to NHS England under the Health and Social Care Information Centre (Transfer of Functions, Abolition and Transitional Provisions) Regulations 2023.

The Health and Social Care Act 2012 ('the Act') gives NHS England statutory powers, under section 259(1)(a), to require data from health or social care bodies, or organisations that provide publicly funded health or adult social care in England, that it considers necessary or expedient to have to carry out its functions under chapter 9 of the Act. This includes where it has been directed to establish an information system by the Secretary of State for Health and Social Care.

The data, as specified by NHS England in the associated Data Provision Notice (DPN), is required to carry out its functions conferred on it by the COVID-19 Public Health Directions 2020. Therefore, organisations that are in scope of the notice are legally required, under section 259(5) of the Act, to provide the data in accordance with the DPN.

The OpenSAFELY COVID-19 service (referred to as the **Service**) provides a secure analytics service for Approved Users (academics, analysts and data scientists) to access GP and NHS England pseudonymised patient data for COVID-19 research, COVID-19 clinical audit, COVID-19 service evaluation and COVID-19 health surveillance purposes (**COVID-19 Purposes**[1]). The Service is currently operated by NHS England in collaboration with the Bennett Institute and The Phoenix Partnership (**TPP**), or Egton Medical Information Systems (**EMIS**) (the **GP System Suppliers**),

This service previously operated under notices issued under regulation 3(4) of the Health Service (Control of Patient Information) regulations 2002 (the COPI regulations). The current notice will expire on the 1st July 2023. As such, and to enable the continuation of the service, the legal basis is now being transferred to the COVID-19 Public Health Direction 2020 issued by the Secretary of State.

[1] The specific terms are defined here: https://www.hra-decisiontools.org.uk/research/docs/DefiningResearchTable_Oct2022.pdf

c. The Service is designed to keep patient data confidential: users write analysis code away from the patient data and test it on dummy data; the Service then automates the running of code (the **Queries**) to generate intermediate pseudonymised patient level datasets (**Intermediate Outputs**) and then final anonymous aggregated outputs (the **Aggregated Outputs**).

The Service uses the OpenSAFELY Platform to run Queries on pseudonymised GP and NHS England patient data which is held within the GP system suppliers' data environments. The list of NHS England pseudonymised datasets held within the Service is set out in Appendix 3.

Approved Users only have access to the Aggregated Outputs generated by their Query. These Aggregated Outputs are released outside the GP System Supplier secure environment only after disclosure controls have been applied within the GP System Supplier secure environment and the results reviewed and cleared by trained output-checkers. All actions by all users in the Service are logged in public, in real-time and all Queries are logged and published. No record level GP or NHS England data leaves the GP system suppliers' environment.

More information about how the Service operates is available on the OpenSAFELY website here: [https://www.opensafely.org/]

As GP Practices are already making data available for this Service, they will not need to instruct their GP System Suppliers to make data available as this is a continuation of an existing service, covering all practices in England that use The Phoenix Partnership (TPP) GP system or the Egton Medical Information Systems (EMIS) GP system.

There are three notable changes to the current arrangement:

1. The legal basis changes from the COPI regulations to the COVID-19 Public Health Directions 2020.

2. Type 1 Opt-outs will be honoured for all newly approved projects under the new legal basis. A fully transparent transition process will be in place for all ongoing projects to enable those to complete without impacting the study results and findings. A list of ongoing projects and their purposes is included in the DPN and within this document as Appendix 3. This approach has been independently reviewed with the Interim Data Advisory Group and the National Data Guardian (NDG) Panel.

3. There is a risk that GP Practices, as Controllers of the GP patient data, may exercise their rights as data controllers and could decide not to comply with the on-going legal obligation to provide the data.

Why is a DPIA required and what is its scope?

The General Data Protection Regulation (**GDPR**) requires a Data Protection Impact Assessment (DPIA) to be completed by a controller where its processing of personal data is considered to be a high risk to the rights and freedoms of individuals. In particular GDPR requires a DPIA to be carried out where there is processing of personal data relating to health on a large scale.

For the GP data, the GP Practices are the Controllers of the data within the clinical systems, with GP System Suppliers (GPSS) as Processor. The data is then processed (by the GPSS) under instruction of the GP Practice as Controller, to make a de-identified and pseudonymised copy of only the coded patient data that continues to be stored within the GPSS secure boundary (previously referred to as OpenSAFELY-EMIS/-TPP). The GP Practice remains Controller of the data at this point.

Separately, NHS England ingests other (non-GP controlled) pseudonymised data (for which NHSE is the Controller) within the secure boundaries of the GPSS. The GPSS are acting as Processor to NHS England for this activity. Details on those data sets can be found in Appendix 1.

NHS England then sends a Query to the GPSS, executed via the OpenSAFELY system, and the GPSS runs that query on both the GP-Controlled data and also the NHSE-Controlled data and returns an NHSE-Controlled Output to NHSE (this Output remains within the secure boundaries of the GPSS, who act as Processors for NHSE for this activity).

The collection and processing of this data by NHS England requires a DPIA.

The processing of the pseudonymised data in each data set is carried out by the NHS England OpenSAFELY Service to support **COVID-19 Purposes**

Therefore, this document has been prepared by NHS England, as a DPIA to satisfy its own compliance requirements as a Controller of personal data, as described above.

This document should be read in conjunction with the DPIA Guidance and DPIA Screening Questionnaire (see below):

DPIA Screening Questions:

Will the processing involve a large amount of personal data (including pseudonymised personal data) and affect a large number of data subjects?	YES. The process will involve processing pseudonymised GP data (for all patients in England whose GP Practice uses TPP/EMIS as a System Supplier) with a range of other related pseudonymised datasets (see Appendix 1)
Will the project involve the use of a new technology(ies)?	NO.
Is there the risk that the processing may give rise to discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy (e.g., health records), unauthorised reversal of pseudonymisation[3], or any other significant economic or social disadvantage?	No, because unauthorised reversal of pseudonymisation is not possible by the research analysts. Only NHS England, the GP System Suppliers (EMIS and TPP) and the external data providers technical staff have access to the encryption keys. No system administrators and platform developers have access to the encryptions keys. The access to the keys is rigorously controlled by NHS England to only those who need it. Furthermore, once TPP prepare the datasets for analysis, they replace the pseudonym with another pseudonym known only to them (for which only TPP hold the matching table for). EMIS retains the original pseudonym. This significantly

	reduces the risk of re-identification and protects the patients.
Is there the risk that data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data?	NO.
Will the processing of personal data occur without informing the individual of the processing?	YES. Controllers of personal data have obligations to make patients aware of any processing of their data. NHS England is working with the GP profession to ensure there is a coherent approach to transparency for the public, but patients will not be aware of specific processing activity on their own data as the information is pseudonymised and de-identified. Transparency materials will be available on the former NHS Digital website and will be provided to practices to publish for their patients.
Will there be processing of genetic data, data concerning health or data concerning sex life?	YES - health data; no genetic data (we believe the answer is “no”: we may have access to genetic diagnostic codes, for example stating that someone has a disease with a genetic component (e.g. “Down's Syndrome”) but we will not currently have access to individuals’ complex genome or actual gene sequencing information as these are not typically stored as structured or coded data in GP records; there will be coded data items on sexual orientation and other sexually transmitted diseases shared by the patient or laboratories, or from correspondence received from sexual health services which are subsequently coded into the GP record; but at no point will any sexual health clinical data be automatically included unless it is recorded directly by the GP.

Are the data to be processed revealing racial or ethnic origin, biometric data, political opinions, religion or philosophical beliefs, or trade union membership?	YES - some of these are recorded by GPs in the patient record. Ethnic origin data will be required for analysis purposes and will already be present in the GP records and in other health data sources linked via OpenSAFELY and controlled by NHS England such as SUS.
Will there be processing of data concerning criminal convictions and offences or related security measures?	NO.
Will personal data of vulnerable natural persons, in particular of children, be processed?	YES.
Will personal aspects be evaluated, in particular analysing or predicting aspects concerning performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, in order to create or use personal profiles?	NO.
Will the project include a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person (e.g., a recruitment aptitude test which uses pre-programmed algorithms and criteria)?	NO.
Will there be a systematic monitoring of a publicly accessible area on a large scale (e.g., CCTV)?	NO.
Will the processing include any data matching e.g., the combining, comparing, or matching personal data obtained from multiple sources?	YES. Data from the datasets listed in Appendix 1 will be linkable to pseudonymised GP records held in TPP and EMIS.

Will the processing include any tracking e.g., processing which involves tracking an individual's geolocation or behaviour, including but not limited to the online environment?	NO.
Will the processing include any denial of service e.g., decisions about an individual's access to a product, service, opportunity, or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data?	NO.

If the answer to any of the above questions is Y, please complete the rest of the form. If all of the screening questions are answered N, the local IG team must still sign off the DPIA.

4. Consultation with Stakeholders

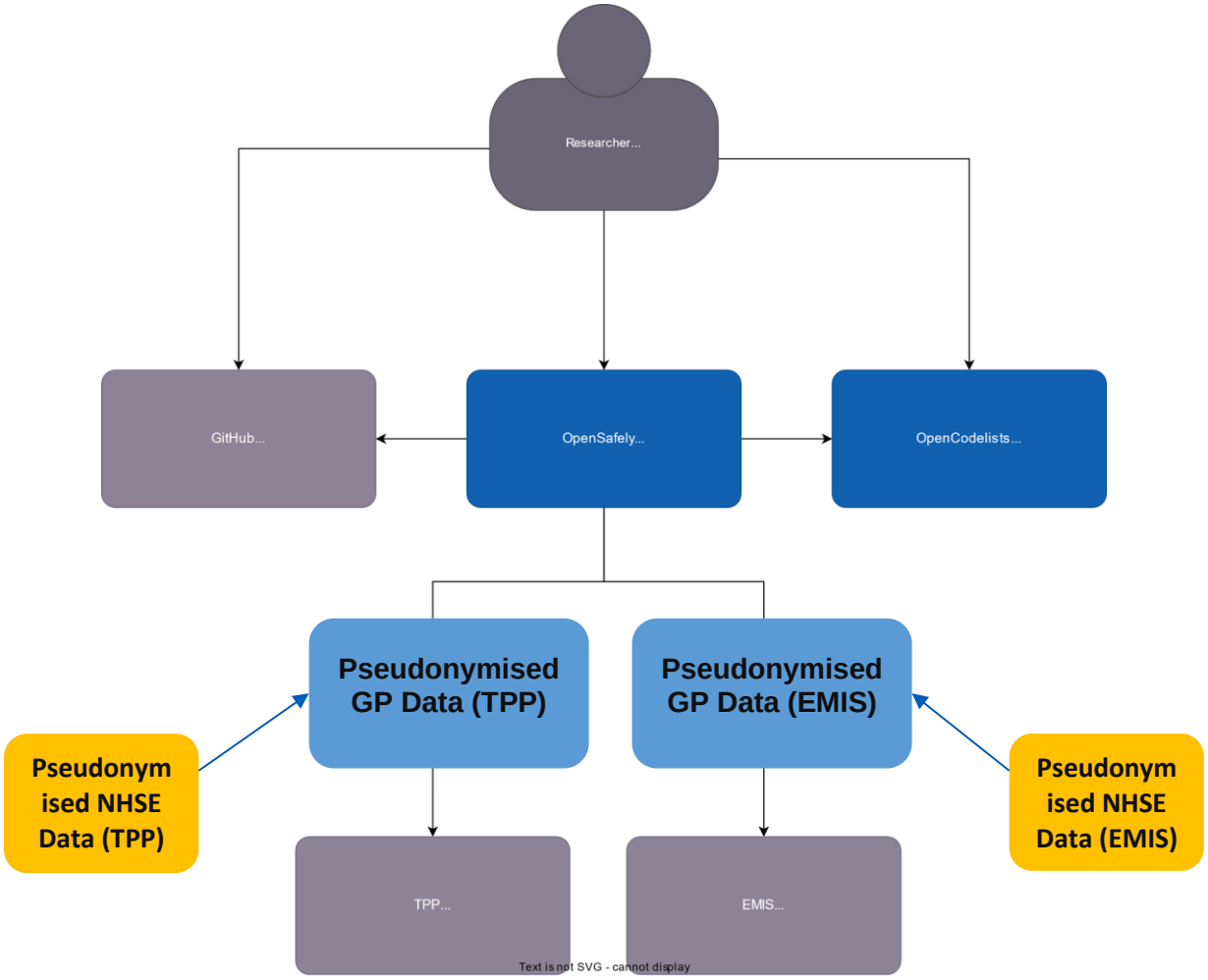
NHS England has, as required under section 258 of the Health and Social Care Act 2012, consulted with the following persons and groups:

- Department of Health and Social Care, as directing organisation
- The British Medical Association (BMA)
- The Royal College of General Practitioners (RCGP)
- The BMA and RCGP Joint GP IT Committee
- The National Data Guardian for Health and Social Care (NDG)
- Information Commissioners Office (ICO)
- Interim Data Advisory Group which replaced Independent Group Advising on the Release of Data (IGARD) (and is in place until the formal terms of reference for the Advisory Group on Data (AGD) are formalised)
- OpenSAFELY Oversight Board
- Representatives of patients including Understanding Patient Data, Healthwatch England, Health Data Research UK patient and public involvement and engagement group, Use My Data, and a small sample of GP patient participation groups via the monthly GP Data Patient and Public Engagement and Communications Advisory Panel.
- GP System Suppliers (EMIS, TPP)
- The Data Alliance Partnership Board (DAPB), which includes representatives from the Department of Health and Social Care, The National Institute for Health and Care Excellence, NHS England

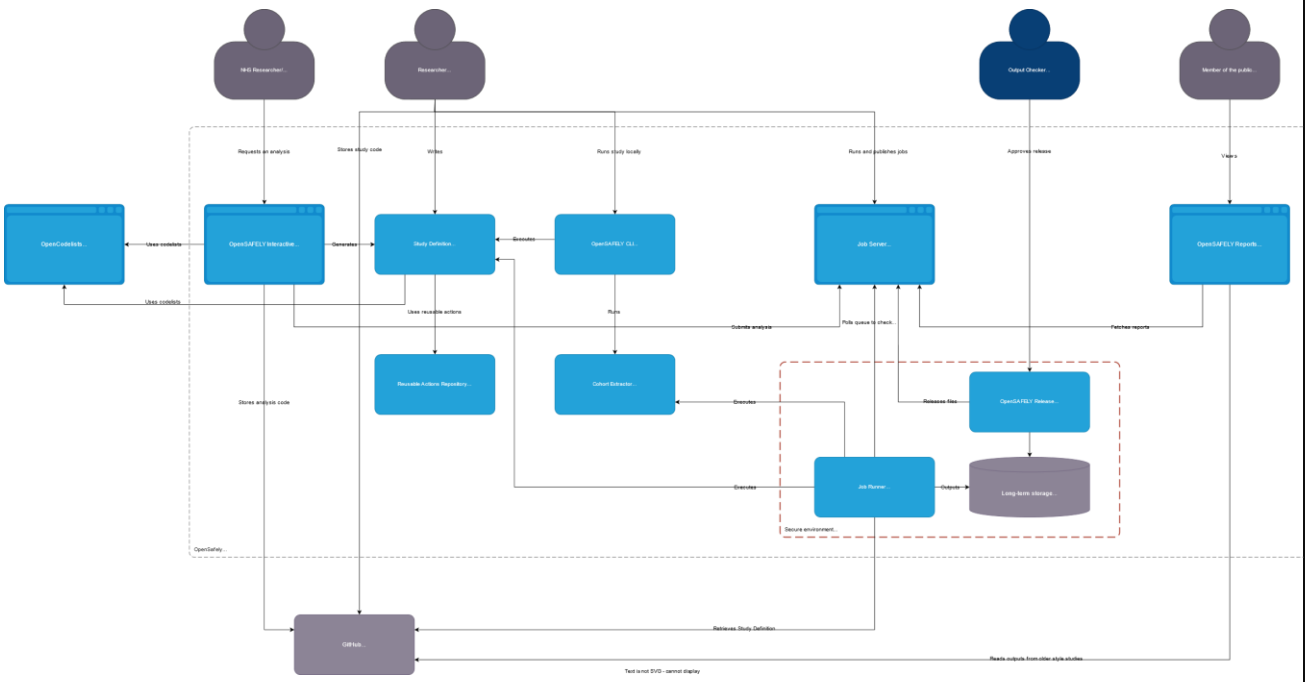
5. Data Flow Diagram

As per the details on the OpenSAFELY website, the data flow diagram is as follows:

High-level diagram:



Detailed Technical Architecture:



The Service has been built with a focus on patient privacy and protections and, as such aims to mitigate risk through the use of its tools and design.

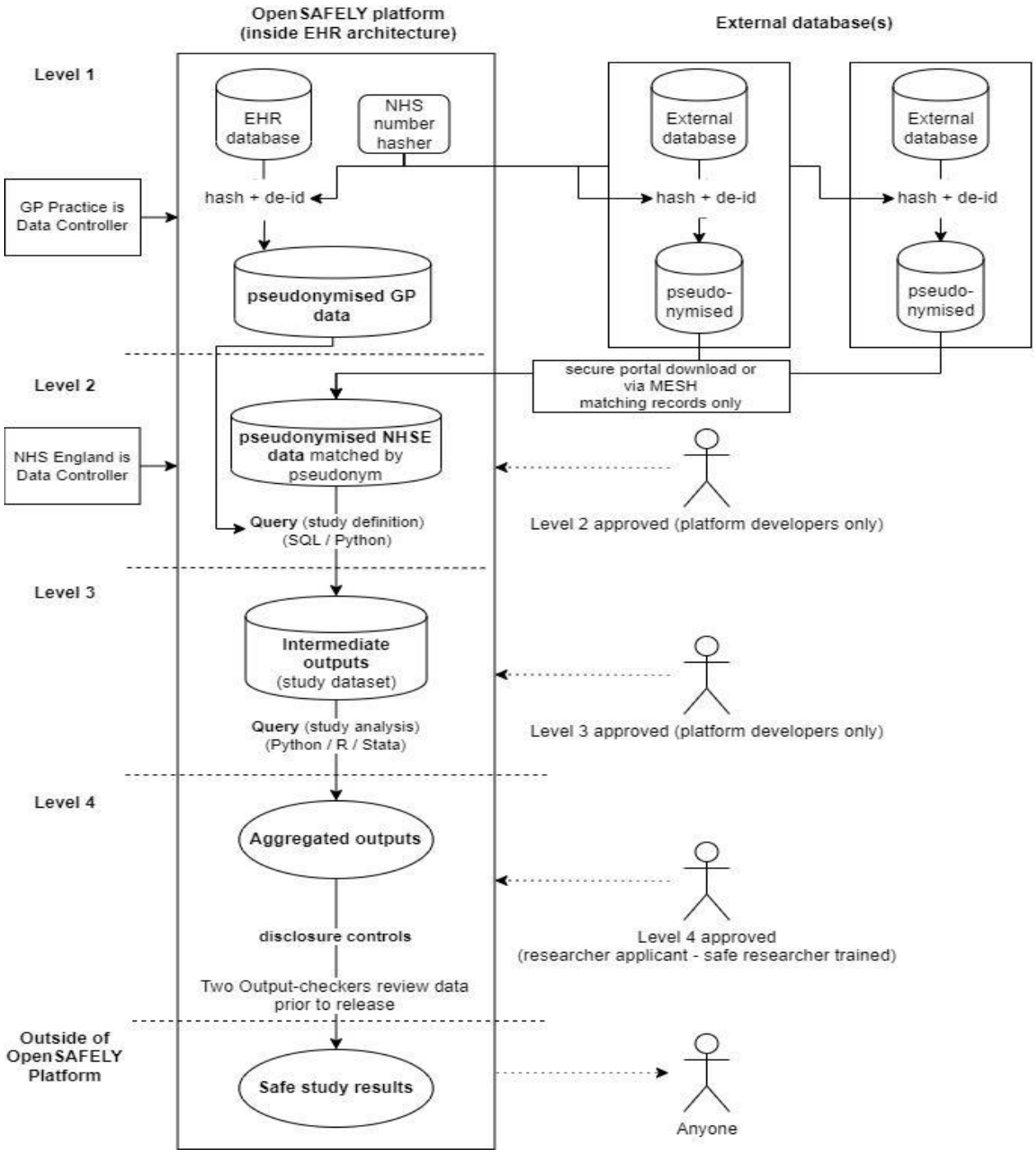
“Pseudonymisation” is a widely used process for protecting patients’ privacy whereby explicit identifiers such as names, addresses, and dates of birth are removed from patients’ medical records before they are used or shared. Pseudonymised data can also be referred to as de-identified data but is treated as identifiable as it is technically possible to re-identify patients from pseudonymised data, without other controls and measures in place.

The pseudonymised datasets held within TPP and EMIS are then analysed using the OpenSAFELY platform. This platform allows users to design their query and run their code against randomly generated dummy data, before running the actual query as an automated request on the pseudonymised data within the service. This eradicates the need for users to develop their analyses by interacting directly with real pseudonymised patient data, and thus eradicates a key historic source of risk for projects where data is accessed at very large population scales.

The event-level datasets (both GP-Controlled and NHSE-Controlled) are queried by the service. Intermediate Outputs (patient-level, under NHSE Control) are the results from the Queries. These Intermediate Outputs are then subject to statistical analysis - again *without the users having sight of this data* - to generate Aggregated Outputs. A limited number of Approved Users can access the Aggregated Outputs within the Service (subject to a Data Access Agreement). The Aggregated Outputs are considered as anonymous at this point; however, even then they cannot be released/removed from the system or linked with other data. Prior to release from the system, there is a disclosure control process carried out to ensure the Aggregated Outputs will remain anonymous outside the boundaries of the secure service. The disclosure control process involves the Approved Users applying disclosure controls to the results they seek to release; these results are then reviewed and cleared for release by two trained output-checkers.

As per the below diagram, the pseudonymised GP data remains under the control of GP practices sitting inside Level 1. The Service automates the running of code (the Queries) against pseudonymised GP data (level 1) and pseudonymised NHS England data (level 2), to generate intermediate pseudonymised patient level datasets (Intermediate Outputs) in Level 3, and then final anonymous aggregated outputs (the Aggregated Outputs) in Level 4.

Detailed flow of data by level:



Data Inflows Summary table:

Appendix 1

Data Outflows:

All outputs are anonymised/aggregated.

6. Purpose of the processing

The Secretary of State has directed NHS England to collect, process and analyse data in connection with COVID-19 to support the Secretary of State's response to COVID-19 and support various **COVID-19 purposes** set out in the COVID-19 Public Health Directions 2020, 17 March 2020 (as amended). This enables NHS England to collect data, analyse and link the data for **COVID-19 purposes** with other data held within the secure environments within TPP and EMIS with access via the NHS England OpenSAFELY service.

Since the expiration of the General COPI Notice, the NHS England OpenSAFELY service in EMIS and TPP (previously OpenSAFELY-EMIS/-TPP) has been operating under notices issued by the Secretary of State for Health and Social Care under regulation 3(4) of the COPI regulations . The current notice is due to end on the 1 July 2023 at which point the legal basis for the NHS England OpenSAFELY service will transfer to the COVID-19 Public Health Directions 2020. This DPIA, published in association with the *NHS England OpenSAFELY (OS) COVID-19 Service DPN* ensures that, following the transfer of the legal basis, GP practices will continue to be required to pseudonymise their data held within EMIS and TPP and provide the service with access to run approved queries against these pseudonymised patient records.

The NHS England OpenSAFELY service is used to provide access to these stores of pseudonymised personal data to support approved projects related to COVID-19 infection. This service provides users with the ability to remotely analyse and gain insight from linkable GP patient data within the GP System Supplier boundary for approved COVID-19 research, COVID-19 clinical audit, COVID-19 service evaluation and COVID-19 health surveillance purposes. The continuation of this collection and the subsequent analysis is necessary to prevent disruption of this COVID-19 research.

The NHS England OpenSAFELY service will be used to link de-identified record level data from pseudonymised healthcare datasets, with pseudonymised patient data held on systems operated by TPP and EMIS (GP System Suppliers, GPSS), to support the **COVID-19 purposes** (See Appendix 1 for full list of external datasets). The analysis is conducted by NHS England analysts (including those under honorary contracts) and/or 3rd parties who have data access agreements with NHS England.

Further details are available here: NHS England » OpenSAFELY – the Coronavirus (COVID-19) Research

1. What is intended to be done with the personal data collected / used / processed / stored during this project?

NHS England will only approve the use of the data for analysis for **COVID-19 purposes**.

Personal data (and data relating to the deceased if they were registered with a GP on 1st January 2009) will be collected and stored in pseudonymised form within TPP and EMIS. Users will not access this data, and most will only be able to view published anonymous, aggregated results with disclosure controls applied (the exceptions to "most" are authorised users per project, who can view aggregated results prior to the application of disclosure controls).

Trained output-checkers who hold NHS England Honorary Contracts or Data Access Agreements) will be able to access de-identified aggregated outputs so that disclosure controls that have been applied by Approved Users are reviewed; once cleared, these anonymous, aggregated, and now checked outputs can be released from the OpenSAFELY platform.

OpenSAFELY is software installed as a safe, productive access layer in front of an existing data warehouse. OpenSAFELY software performs the processing and provides results. These aggregated results are made available for inspection by the authorised per-project user(s) prior to release from the level 4 (Aggregated Output) environment. When a project member wishes to release results more widely (i.e., to share with collaborators for discussion, or to incorporate as part of a manuscript or report), they submit a request for release via the secure output-checking service. Output checking is the final guarantee that outputs are indeed aggregate-only and cannot be used to re-identify patients. Output checked/approved results are published to the "Job Server" dashboard outside the secure environment, from where they can be shared more widely with project members who do not have secure environment access. The results are made publicly available following the final NHSE publication approval step. Once outputs have been released to the Job Server, authorised users can (and are encouraged to) self-publish their output-checked results for public consumption, within a per-project dashboard that shows results alongside analytic code and audit logging.

The workflow with approvals process can found here.

2. How will the personal data be collected (i.e., will it be obtained from the individuals themselves or via a third party)?

Data will be collected from all EMIS and TPP GP Practices in England. The GP practices use GP System Suppliers (EMIS/TPP/others) to store the patients' GP record which contains only data that GPs have already obtained from patients and other third parties, including other healthcare professionals, for the purposes of providing healthcare services to patients. It is not collected directly from the individuals themselves.

Further data sources where NHS England is the owner and Data Controller will be provided and placed in the Level 2 environment in EMIS and TPP alongside the GP data to allow Queries to run against both data sources. These data sources are pseudonymised to protect the identity of the patients but cannot be anonymised at this stage so that linkage can occur. The data collection of these other sources is explained in the relevant DPN, DPIA for that collection and will be available via the NHS Website (where applicable).

Any users using the service will not have direct access to the data. They will develop queries (using dummy data) which are then run against the real data with only the output being made available once disclosure controls have been applied. The published outputs will be anonymous.

3. What will the intended results be, i.e., likely results for a GP Practice, impact (positive and negative, as applicable) on individuals concerned or (where applicable) other parties involved?

General practices have been overwhelmed with requests for access to data held in GP medical records to support the COVID-19 response for vital research purposes. The data held within TPP and EMIS and made available to the service to query, providing NHS England and Approved Users with insight from this data for **COVID-19 Purposes** which significantly reduces the burden on General Practice, enabling General Practitioners to focus on delivering health care and support to patients.

Consequently, users, and those who carry out research, service evaluation, clinical audit, and health surveillance in England for **COVID-19 purposes** will have more timely access to the data they require, assured through a transparent and robust governance process, providing scrutiny and transparency on the use of data about patients. However, it could impact local data sharing agreements between individual GP practices and IT suppliers for other research purposes.

4. **What will be the benefits to the individuals concerned or (where applicable) other parties involved (including GP Practices) and to society?**

NHS England and Approved Users need insight from this vital data for research into COVID-19, to help address the national response to the COVID-19 pandemic and future health pandemics, which will help save lives. **COVID-19 purposes** for which this data may be analysed and used include:

- understanding COVID-19 and risks to public health, trends in COVID-19 and such risks, and controlling and preventing the spread of COVID-19 and such risks
- identifying and understanding information about patients or potential patients with, or at risk of COVID-19.
- understanding information about patient access to health services as a direct or indirect result of COVID-19, and the availability and capacity of those services
- health services research on changes in clinical activity and population impact during the pandemic, as well as consequences of the pandemic; conducting research in relation to COVID-19, including COVID-19 related clinical trials.

Supporting such research via the Service will significantly reduce the burden on General Practice at a time when demand on resources is high, enabling General Practice to focus on delivering health care and support to patients.

7. Description of the Processing

Nature and scope of the processing:

The Service is designed to keep patient data confidential: users write their analysis code (Query) away from the patient data; the service automates the Query being run against the data sets. Only Aggregated Outputs are shared with users. All actions performed within the service are logged in public and in real-time at https://jobs.opensafely.org/ and no record level GP data is shared outside of the GP system environment; disclosivity checks ensure all Aggregated Outputs leaving the platform are non-disclosive and anonymous.

A de-identified and pseudonymised copy of coded GP patient data from the system suppliers EMIS and TPP is held in the Level 1 environment (one for EMIS and one for TPP). This copy of the GP data contains the coded and structured data of all patients in England registered at TPP or EMIS practices. (Some practices use Cegedim as their GP system supplier which accounts for around 1% of GP practices in England (~60 practices as of June 2023). The patients registered at these practices are not included in projects conducted using the service.)

All GP patient data is pseudonymised at source by the relevant supplier before being moved into the Level 1 environment. Further data sources are then ingested into each GP system supplier's secure environment so that linkage can occur and the OpenSAFELY software can perform queries on the data sets. The Services run Queries to generate final Aggregated Outputs (with disclosure controls applied) before any data leaves the platform in EMIS and TPP.

A small number of platform developers (in their role as systems integration on behalf of NHS England) need more direct access to query the pseudonymised data (both GP and non-GP data) inside the secure environment for systems integration and platform development purposes. This access is necessary to ensure the system is operating correctly for users, and controls are built in to ensure the required access is minimised, transparent, and secure. All developer access to platform is logged and on a secure connection; the pseudonymised data developers can query is read-only; their queries are fully logged; there is no facility to enable the extraction of patient or record level data outside of the environment (only Aggregated Outputs). These activities, and the required access levels, are only undertaken by the core OpenSAFELY development team under NHSE honorary contracts and have role-specific security controls to reflect their increased access to data. Further information on the activities covered under these roles and the governing policies are available here: https://docs.opensafely.org/developer-access-policy/#policy-for-opensafely-access-by-platform-developers

Note on planned future changes in scope:

One dataflow is currently in process for ingesting into OpenSAFELY: - King's College London (DSA signed)

Context of the processing:

Outputs will inform NHS and Approved Users on limiting and mitigating the spread of COVID-19; provide evidence around risk factors to inform the provision of care; provide evidence on treatments that may improve outcomes. Identifiable data is not extracted from the service, rather anonymous aggregated results are produced from the processing and provided to users.

8. Describe the legal basis for the processing (collection, analysis, or disclosure) of personal data?

Legal basis for collection, analysis, and dissemination:

NHS England has been directed by the Secretary of State for Health and Social Care under section 254 of the Health and Social Care Act 2012 to establish and operate a system for the collection and analysis of the information specified for this service. The direction is published on the NHS England website: COVID-19 public health NHS England Directions.

This information is required by NHS England under section 259(1)(a) of the Health and Social Care Act 2012.

In line with section 259(5) of the Act, all organisations in scope, in England, must comply with the requirement and provide information to NHS England in the form, manner and period specified in the associated *NHS England OpenSAFELY (OS) COVID-19 Service* Data Provision Notice (DPN).

The above DPN is issued in accordance with the procedure published as part of an NHS England duty under section 259(8).

9. Demonstrate the fairness of the processing

The NHS England OpenSAFELY service works on a direct copy of the coded and structured data in the back end of TPP/EMIS. Specifically, a pseudonymised copy is available for processing in the Level 1 environments in TPP and EMIS. There is no code modification or grouping structures or alteration of the codes that are made available.

Data quality underpins all studies run using NHS data, including all projects run through the NHS England OpenSAFELY service, however, the data quality of specific data sources and variables within those data sources is not well documented. To address this challenge, the NHS England OpenSAFELY service supports short data reports, which provide a carefully documented source of information on data quality that is beneficial to all users of NHS data. These cover four key areas:

- Clinical reports describing how a specific clinical area is recorded in NHS data.
- Demographic reports describing population characteristics and how they can be identified.
- Administrative reports describing patterns in NHS data.
- Methodological reports describing data processing techniques for NHS data.

The diagnosis of COVID-19 infection, prior medical conditions, prescribed medications, allergies, physiological parameters (e.g., BMI, sats, etc), outcome and level of care for COVID, prior investigations such as blood tests and their results, demographic data including age, sex, ethnic origin, and cause of death are all required in order to provide high quality research into COVID-19. This pseudonymised, event level data, is required in

order to support national level risk profiling for long term conditions and/or medications in relation to COVID-19 infection and outcomes.

The data processed is only the data necessary to enable users to understand different risk profiles for people with different medical history and demographic data, in relation to COVID-19.

Protection is applied to the GP data (at this stage under GP Controllership, and therefore outside the scope of this DPIA but included for context) by the removal of intentional identifiers.

Specifically, the following is done by EMIS/TPP in the secure environments in TPP and EMIS:

· Removal of associational identifiers: Mobile phone numbers, email addresses, telephone numbers, hardware and software unique identifiers, IP addresses.

· Removal of transactional unique identifiers: all unique booking reference numbers for appointments, contacts, and referrals.

· Removal of functional unique identifiers: Titles, forenames, middle names, surnames, full dates of birth, full dates of death, house name, house number, street, full postcode.

· Removal of narrative text (commonly referred to as 'free text' data: All free text on patient records is removed. In line with other UK primary care research database permissions, the dosage and quantity fields on prescribed medication are retained, but any script notes are removed.

· Removal of additional unstructured context: scanned images, medical drawings, letters, and all other record attachments.

· Derived data items and removal of exact original values: date of birth (MM/CCYY), partial postcodes at sector level, indices of multiple deprivation, the rurality-urban classification, geographic super-output area codes at each super output area level. Note – for organisations, the only geographic indicators stored are the lower super output areas and / or middle-level super output area code and the Local Authority code and STP/ICS code.

· Pseudonymisation is applied to the remaining data: Generation of strong pseudonyms using industry-standard cryptographic hash techniques with only NHS England approved holders of the keys to prevent re-identification of patients.

10. What steps have you taken to ensure individuals are informed about the ways in which their personal data is being used?

New transparency notices will be developed to communicate the changes to patients and issued to all GP practices with the new DPN. This content will be displayed by GP practices for their patients. This aligns with the NHS England focus on open data principles and the details contained in the General Transparency Notice available on the NHS England website. Transparency information will be provided to GPs for their patients, and information about the data collection and processing will be available on the former NHS Digital website and transferred to the new NHS England website when appropriate.

The existing NHS England privacy notice will be updated in alignment with the changes described in this document. The previous version is available here: https://www.england.nhs.uk/contact-us/privacy-notice/how-we-use-your-information/COVID-19 -response/coronavirus-COVID-19-research-platform/

In addition, NHS England has communicated to the public on the use of OpenSAFELY in the following two articles posted on the NHS England website:

OpenSAFELY: secure access to data to deepen our understanding of COVID-19 - Improving care through research and innovation - NHS Transformation Directorate (england.nhs.uk)

NHS England » OpenSAFELY – the Coronavirus (COVID-19) Research Platform

Separately, the OpenSAFELY website describes the software and its operation along with hosting a publicly available list of data sources: Data Sources - OpenSAFELY documentation. All research projects using OpenSAFELY are made public here: https://www.opensafely.org/approved-projects/.

OpenSAFELY has also worked with Citizens Juries to validate the design of the software and its operation: Data Sharing in a Pandemic: Three Citizens' Juries Juries' Report (nihr.ac.uk)

11. Is it necessary to collect and process all data items?

Data Categories [*Information relating to the individual's*]	**YES/ NO**	**Justify** *[there must be justification for processing the data items. Consider which items you could remove, without compromising the purpose for processing]*
Personal Data		
Name	NO	
Address	NO	
Postcode	YES	Partial postcodes are processed. It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
DOB	YES	Partial DOB in format MM-YYYY. It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality. This variable informs calculation of age (see below).
Age	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Sex	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the

Data Categories [*Information relating to the individual's*]	**YES/ NO**	**Justify** *[there must be justification for processing the data items. Consider which items you could remove, without compromising the purpose for processing]*
		spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Marital Status	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Gender	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Living Habits	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Professional Training / Awards / Education	YES	Such as occupation codes. It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Income / Financial / Tax situation / Financial affairs	NO	
Email Address	NO	
Physical Description	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
General Identifier e.g., NHS No	YES	NHS Numbers are pseudonymised to protect the identity of patients, but it is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Home Phone Number	NO	
Online Identifier e.g., IP Address/Event Logs	NO	
Website Cookies	NO	
Mobile Phone / Device No / IMEI No	NO	
Location Data (Travel / GPS / GSM Data)	NO	
Device MAC Address (Wireless Network Interface)	NO	
Banking information e.g., account number,	NO	

Data Categories [*Information relating to the individual's*]	**YES/ NO**	**Justify** *[there must be justification for processing the data items. Consider which items you could remove, without compromising the purpose for processing]*
sort code, card information		
Criminal convictions / alleged offences / outcomes / proceedings / sentences	YES	as SNOMED codes, not free text, where a GP has (infrequently) chosen to capture such coded information relating to this as part of the health record.
Special Category Data		
Physical / Mental Health or Condition	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Sexual Life / Orientation	Partial	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality. Data is collected and processed except: a. SNOMED Refset for 'General Practice summary data sharing exclusion for gender related issues' 999004371000000109 b. SNOMED Refset for 'General Practice summary data sharing exclusion for assisted fertility' 999004351000000100 c. SNOMED Refset for 'General Practice summary data sharing exclusion for termination of pregnancy' 999004361000000107 d. All children of the SNOMED code 118199002 'Finding related to sexuality and sexual activity'
Religion or Other Beliefs	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Trade Union membership	YES	as SNOMED codes, not free text, where a GP has (infrequently) chosen to capture such coded information relating to this as part of the health record.
Racial / Ethnic Origin	YES	It is necessary to identify risk factors, treatment options that might improve outcomes and approaches to reducing the spread, treatment, and prevention of the disease. The public interest served by identifying risk factors and slowing the spread of disease outweigh the duty of confidentiality.
Biometric Data (Fingerprints / Facial Recognition)	NO	
Genetic Data	NO	We believe the answer is "no": we may have access to genetic diagnostic codes, for example stating that someone has a disease with a genetic component (e.g., "Down's Syndrome") but we do not have access to individuals' complex genome or actual gene sequencing information as these are not typically stored as structured or coded data in GP records.

Publicly available list of data sources: Data Sources - OpenSAFELY documentation

12. Describe if personal datasets are to be matched, combined, or linked with other datasets (internally or for external customers)

NHS England uses two methods for matching the pseudonym. Matching by external data provider: EMIS or TPP sends a file of all their patient pseudonyms to the external data providers; only patient records with a match in the external databases are transferred back into the Level 2 environments in EMIS and TPP. This is the preferred mechanism as it minimises unnecessary data flow.

Matching by EMIS and TPP: On some occasions (due to technical constraints of the external database provider or when the external provider's population is much smaller, relative to that of the data sets), the pseudonym matching occurs inside the Level 2 environment in EMIS and TPP. This involves a two-step process:

- The external data provider makes available one file of their full pseudonym list to EMIS and TPP which in turn establish a list of matches of the pseudonyms they have. Each GP System Supplier sends back a list of these matched pseudonyms to the external data provider in a file.

- The external data provider then prepares separate files for each of the data sets according to the matched list they were provided with. The two separate files produced by the external data provider, one for EMIS and one for TPP, containing the additional requested and approved dataset for only the matched pseudonyms provided in step one. This data is then added to the Level 2 environment in EMIS and TPP.

For record linkage, the pseudonyms will be produced by the GP System Suppliers who will hold the keys. This pseudonym with additional data items provides significant confidence in matching accuracy, ensuring that the right record is linked to the right primary care record. Research models will be automatically executed as required for analysis. GP patient data held by TPP and EMIS is incrementally updated to reflect changes made by clinicians in England GP practices.

The patient data processed for this purpose includes:

- Demographic information (age, sex, area of residence, ethnicity),
- Clinical information pertaining to COVID-19 care and outcomes,
- Clinical information pertaining to wider health conditions, medication, allergies, physiological (e.g., BMI), prior blood tests and other investigation results, and other recent medical history (e.g., smoking status).

The patient population made available within the NHS COVID-19 OpenSAFELY Databases in EMIS and TPP for analysis is defined as all patients EXCEPT:

- Patients who have no period of registration (in a TPP or EMIS practice) after 1 January 2009 or who died before 1 January 2009 (year of Swine Flu Pandemic)

This means that for each system supplier if there has been any period of registration after 1 January 2009 then all the patient's data from the last practice with that system supplier is available to the Service for analysis, irrespective of the patient's current registration status.

Patients' Type 1 Opt-outs will be upheld by default by the system with only previously approved projects retaining access to Type 1 data as part of a transitionary arrangement

to the new legal basis as agreed with Interim Data Advisory Group and the NDG Panel. The list of these projects is included in Appendix 3.

Patient data made available to NHS England via the data stores in EMIS and TPP is pseudonymised and de-identified at source.

For this patient population, all coded patient data will be made available for processing, except the following codes in these reference sets:

- SNOMED Refset for 'General Practice summary data sharing exclusion for gender related issues' 999004371000000109
- SNOMED Refset for 'General Practice summary data sharing exclusion for assisted fertility' 999004351000000100
- SNOMED Refset for 'General Practice summary data sharing exclusion for termination of pregnancy' 999004361000000107
- All children codes of the SNOMED code 118199002 'Finding related to sexuality and sexual activity'.

Summary of the specific data sets that are matched, combined, or linked using the NHS England OpenSAFELY service are described above in Section 3 on the data flows. All linking happens within the NHS England OpenSAFELY service and no linkage can occur outside the data sets, as outputs are in anonymised, aggregated form.

The transfer mechanism varies on the data provider and outlined in Appendix 1.

Data sharing agreements are in place to support the current process of linking GP data and non-GP data flowing to both of the data stores (in EMIS and TPP). The data store in TPP contains GP data from TPP, linked to (for example) ONS Deaths; and the data store in EMIS contains GP data from EMIS, linked to the same datasets. Neither data store contains GP data from the other; and the only linked data they contain is that relating to their own populations.

The process of ingesting pseudonymised NHSE data may be altered in the future according to the wishes of NHS England.

All externally linked datasets currently flow into Level 2 environments in TPP and EMIS.

See the Data Provision Notice for more information.

13. Describe if the personal data is to be shared with other organisations and the arrangements you have in place

Personal data is not shared with other organisations. Outputs are in aggregated and anonymous form, with disclosure controls in place according to the Code of Practice for Statistics. This means that while the service processes pseudonymised data, the research analysts submit a query (that was prepared on randomly generated dummy data) which is then run against the pseudonymised personal data held securely within TPP and EMIS.

This returns an aggregated and anonymous output. Therefore no “personal” data is shared or disseminated by design. The research analysts cannot view personal data.

14. How long will the personal data be retained?

The retention period is in line with the NHS Digital Records Management Policy and the NHS Records Management Code. The pseudonymised and de-identified patient level summary data will be retained for at least 2 years for verification of analyses and for audit proposes. All data will be held according to GDPR regulations. Furthermore the NHS Records and Document Management Policy and Corporate Retention and Disposal Framework: Implementation Process will be adhered to.

Any extension to the direction that would allow data to continue to be retained will be published here.

As the NHS data stores are held with the data processors TPP and EMIS, the data will be held for verification of findings and audit purposes and can be deleted once outside the retention periods, however the data cannot be transferred to other NHSE systems (as per operating requirements set out by the Secretary of State) .

If a patient applies a type one objection, their data (including historical data) will be removed from the data set within one week. The type one objection would be applied to all pseudonymised data sources held within TPP and EMIS and available to be queried via the service (e.g., Level 1 and Level 2 environments). Where queries have already been run prior to the Type 1 objection being applied and Intermediate Outputs have been generated (e.g., Level 3 onwards), these data sets would continue to hold the Type 1 data to ensure the integrity of the research study that ran the query

15. Where you are collecting personal data from the individual, describe how you will ensure it is accurate and if necessary, kept up to date

No personal data is collected from the individual.

16. How are individuals made aware of their rights and what processes do you have in place to manage such requests?

Individuals (data subjects) have the following rights under GDPR:

- The right to be informed – Fair Processing information and Transparency Notices for NHS England and GPs will be developed by NHS England and made available to GP Practices. These will be aligned accordingly to reflect the changes described in the DPIA and DPN. The existing privacy notice on the NHS England website will also be updated

- The right of access - An explanation about how an individual can request a copy of information that NHS England holds is published at: https://digital.nhs.uk/article/6851/How-to-make-a-subject-access-request. Due to the pseudonymisation of the data held within EMIS and TPP, it would require a re-identification of the records to provide details of the data used in the service. Consequently, NHS England will provide information of the source data that NHS England holds on an individual.

NHS England has established processes for handling Subject Access Requests through the Information Governance Team (IG Helpline Service). NHS England has established that a patient record can be extracted, and the explanatory textual description supplied for each of the SNOMED codes. Any patient can also request to see part of their medical records from the Practice either through the GP Practice system supplier or via the NHS App. For data originating from non-NHSE sources, patients can make their requests directly with the source providers (as noted in Appendix 1). Lastly, patients can request access to information in their health records by making a request to the GP Practice under the Access to Health Records Act and by making a Subject Access Request. Information about how to make these requests should be available on GP Practice websites. To minimise the burden on GPs at this time patients are encouraged to register and use NHS App services which include access to medical records.

- The right to rectification - The right for individuals to have inaccurate personal data rectified, or completed if it is incomplete, will be upheld when such a request is received. Patients will need to contact the Practice to ensure that inaccurate data held on GP Practice IT systems is amended. As the data is supplied by EMIS and TPP with a rebuild of the Data sets (currently set at minimum weekly (TPP) or monthly (EMIS), any updates to the patient record will be applied timely via these processes. The data held can receive real time updates on request (daily, or weekly) where required.

- The right to erasure – an individual has the right to request erasure, however NHS England has the right to retain the data where necessary for legal purposes, such as defence of a legal claim and for audit purposes. Data will not be kept longer than necessary in accordance with the NHS Records Management Code of Practice.

.• The right to restrict processing - Where an individual contests the accuracy of their personal data NHS England will consider the request. An individual can also restrict processing by applying a Type One Opt-Out through their GP.

- The right to complain - Individuals who believe that their data is not being processed in accordance with the law can contact NHS England's Data Protection Officer (DPO) regarding NHS England data processing activity and the DPO of their general practice regarding GP data processing activity. Details for the NHS England DPO are contained on the NHS England website in the General Transparency Notice. If this fails, they can complain to the Information Commissioner's Office (ICO).

- The right to data portability- is not applicable to this processing because under article 20 (3) the processing is being carried out in the exercise of official authority vested in the controller under Article 6(1)(c) legal obligation under the COVID-19 Directions or under Article 6(1)(e) public task.

- The right to object – is not applicable to this processing because under article 20 (3) the processing is being carried out in the exercise of official authority vested in the

controller under Article 6(1)(c) legal obligation under the COVID-19 Directions or under Article 6(1)(e) public task.

- Patients that have registered a Type 1 objection with the GP Practice will not have their data shared with the NHS England OpenSAFELY service (or any other organisation outside of their GP practice for purposes other than Direct Care).

- Patients who have registered a National Data Opt-out will have their GP data shared with NHS England. The National Data Opt-out does not apply because that data is required under section 259. No identifiable or patient level data is disseminated to 3rd parties.

- Rights in relation to automated decision making and profiling - no automated decision making, or profiling is intended to take place as part of this processing. However, application of this right will be considered as part of any internal application for access to analyse the data by the NHS England Information Governance Team and by the Advisory Group for Data (AGD) on any application to access and use the data.

In addition, all research projects using OpenSAFELY software are made public here: https://www.opensafely.org/approved-projects/.

17. What technical and organisational controls for "information security" have been put in place?

The governance and service model includes:

1. Application stage – applications are assessed to ensure that:
    - The application is COVID-19 related
    - The data necessary to support the purpose of the application is available in the system.
2. Technical controls – Data is de-identified and pseudonymised at source; linkage of datasets is managed inside by the GP system suppliers (EMIS and TPP). No system administrators or platform developers have access to the pseudonymisation key. No event level or patient level data leaves the service (i.e., the secure network boundaries of EMIS or TPP). Users must specify up front the code they are using to analyse the patient data: explicitly writing studies as "analysis code" means that it is now possible for any interested party to check exactly how patient data was processed, and to assess if such processing is in line with the approved project purpose; the platform principles also require that all analysis code is made public: "**It is accepted that some code may remain private while an analysis is in development. However, all code is published when the results of the analysis are shared (or, for non-complete projects, as soon as possible, usually at the point of their cessation, and no later than 12 months after any code has been executed against the raw patient data)". System administrators can control when analysis code is made public to maintain our transparency principles.** Users cannot access the event level, or patient level data directly; they access the aggregated outputs (which are effectively anonymous) of their analyses over a secure encrypted connection. Access to the aggregated outputs is via a secure encrypted connection, unique to each user, with all access audited. In addition, all researcher actions in

the Service are logged in public, in real-time and all Queries are logged and published (https://jobs.opensafely.org/).

3. People controls – All users must pass safe researcher training (such as that provided by the ONS or UK Data Service) to have access to the secure environment (Level 4) hosting the aggregated outputs. All users who write study code or access the Level 4 environment, and all output checkers, must sign a Data Access Agreement with NHS England. All new users have a 'co-pilot' who is an expert user providing appropriate training and support to ensure safe practice is followed and plausible results are generated.
4. Output checking – Users apply disclosure controls to aggregated outputs they request for release from the secure environment. All outputs are then checked to ensure that they are non-disclosive and safe to release. Each output is independently verified by two output checkers who have been suitably trained.
5. Transparency – all code run using the system is published, as are the research papers that result from the analysis. All accepted applications will also be published.

The two NHS owned data stores in EMIS and TPP are based on AWS Cloud data centres, which are certified for compliance with ISO/IEC 27001:2013 (IT - security techniques - information security management systems), 27017:2015 (IT - security techniques – code of practice for information security controls based on ISO/IEC 27002 for cloud services), 27018:2019 (IT - security techniques - code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors) and ISO/IEC 9001:2015 (quality management systems). Certification details are available at https://aws.amazon.com/compliance/iso-certified/.

Data in transit will be protected by standard security policies available for AWS Elastic load balancers and application load balancers (https://docs.aws.amazon.com/elasticloadbalancing/latest/network/create-tls-listener.html)

<u>Pseudonymisation, de-identification and minimisation at source (Process)</u>

No free text information or directly identifiable patient information is ever transferred; redundant data is pruned (for example, the datasets available to NHS England OpenSAFELY service do not contain the full data schema for SUS data, but a subset); and data about patient geographic locations is always rounded up to the least disclosive level that is still analytically useful (typically, MSOA level for patients, and STP level or LA code for organisations). The level of data available is all publicly documented. Identifiable patient information within EMIS and TPP (and other data provider organisations) is de-identified according to the processes outlined in this DPIA. In all cases, a pseudonym is created with the patient's NHS number in combination with a salt using the SHA2 512 cryptographic hash technique. The salt is shared by TPP/EMIS with the external data provider: [Redacted]

There is no directly identifiable data being processed under the scope of this DPIA. It is acknowledged that the sources of the data being provided (e.g., GP Practices) have access to the identifying data, but that is outside of the scope of this DPIA.

The OpenSAFELY analytics tool operates a tiered security model, described here: https://docs.opensafely.org/security-levels/.

Identifiable data is accessible only in Level 1: this is not available to any users or developers. Level 1 is the data held within the direct care servers of EMIS and TPP, who operate as data processors for GP practices, or the external source data provider environments. The pseudonymised GP data remains under the control of GP practices

inside Level 1. The Service automates the running of code (the **Queries**) against pseudonymised GP data (level 1) and pseudonymised NHS England data (level 2), to generate intermediate pseudonymised patient level datasets (**Intermediate Outputs**) in Level 3, and then final anonymous aggregated outputs (the **Aggregated Outputs**) in Level 4 (See Data Flow Diagram).

Pseudonymised data is accessed via a secure environment provided by the vendor, within their main warehousing infrastructure. [Redacted]. Only aggregated data (after disclosure controls have been applied, as part of the Five Safes Framework) leave the secure environments within EMIS and TPP.

18. In which country/territory will personal data be stored or processed?

The United Kingdom.

The geographical location of the data once transferred to TPP and EMIS is within the UK jurisdiction in their secure environments, which are hosted in the AWS cloud within the UK.

All NHS England staff and contractors accessing the data store will be doing so from within the UK. The OpenSAFELY software platform was built by the Bennett Institute and operates from the UK.

The Data Sharing Agreements restrict the use to defined territories. The application process includes assessing the recipient's legal basis for processing data within these territories, and where such the territory of use is outside the UK. This is reviewed and approved by NHS England according to GDPR and the NHS Records Management Policy.

19. Does the National Data Opt Out apply to the processing?

No. The National Data Opt Out (NDOO) applies to the dissemination of identifiable patient data. OpenSAFELY software ensures that only aggregate, anonymous data can be disseminated using the service. Consequently, the NDOO does not apply to the processing via the service in EMIS and TPP secure environments.

20. Identify and assess risks

Consider the potential impact of your processing and the potential harm or damage that it might cause to individuals whether physical, emotional, moral, material, or non-material e.g., inability to exercise rights; discrimination; loss of confidentiality; re-identification of pseudonymised data, etc.

You can also use this section to detail any risks you have in complying with data protection law and any resulting corporate risks e.g., impact of regulatory action; reputational damage; loss of public trust, etc.

Describe source of the risk and nature of potential impact on individuals	**Likelihood of harm** (Remote; reasonable possibility or more likely than not)	**Severity of impact** (Minimal impact; some impact; or serious harm	**Overall risk rating** (Low; medium; or high)
There is a risk that GPs are not sufficiently aware of the service as it was created during the pandemic with limited communications, and communications distributed alongside the new DPN, which could lead to concern and challenge	Reasonable possibility	Minimal impact	Medium
There is a risk that GP Practices, as Controllers of the GP patient data, may exercise their data protection responsibilities and could decide not to comply with the on-going legal obligation to provide the data, which could lead to a reduction of data available for the service and impact research effectiveness	Remote	Minimal impact	Low

There is a risk that the current data controllership position as it relates to developer access specifically for necessary systems integration is misunderstood, which creates concern, direct challenge, or negative public responses	Remote	Minimal impact	Low
There is a risk that patient confidence regarding NHSE's use of data is impacted due to a lack of historic or current patient awareness regarding the service and the data it uses, which could lead to negative sentiment and increased opt-outs	More likely than not	Minimal impact	Medium
There is a risk that the data held within the system could be re-identified or used to identify patients, which could lead to a breach of GDPR and legal challenge	Remote	Some impact	Low
There is a risk of cybersecurity threats to the service, which could be used to enable access, use, or extraction of data from the service	Remote	Serious harm	Low
There is a risk that research is requested and approved that would be deemed as unethical, which could undermine public trust in the service.	Remote	Some impact	Low

Further Risks are listed in Appendix 2.

20.1. Measures to mitigate (treat) risks

Against each risk you have identified, record the options/controls you have put in place to mitigate the risk and what impact this has had on the risk. Make an assessment as to the residual risk.

Also indicate who has approved the measure and confirm that responsibility and timescales for completion have been integrated back into the project plan.

Risk	**Options to mitigate (treat) the risk**	**Effect on risk** (Tolerate / Terminate / Treat Transfer)	**Residual risk** (Low / Medium / High)	**Measure approved** (Name and Date)	**Actions integrated back into project plan** (Date and responsibility for completion)
There is a risk that GPs are not sufficiently aware of the service as it was created during the pandemic with limited communications, and communications distributed alongside the new DPN, which could lead to concern and challenge	- GP Profession proactive support and comms to be distributed ahead of DPN issuance - System suppliers provided with suitable information to undertake comms through their channels if required - NHS England to distribute comms and patient transparency materials to GPs ahead of DPN issuance	Treat	Low	Agreed between NHS England, OpenSAFELY, and GP Profession	26/05/2023 Redacted
There is a risk that GP Practices, as Controllers of the GP patient data, may exercise their rights as data controllers and could decide not to comply with the on-going legal obligation to provide the data, which could lead to a reduction of data available for the service and impact research effectiveness	- GP Profession proactive support and comms to be distributed ahead of DPN issuance - System suppliers provided with suitable information to undertake comms through their channels if required - NHS England to distribute comms and patient transparency materials to GPs ahead of DPN issuance	Treat	Low	Agreed between NHS England, OpenSAFELY, and GP Profession	26/05/2023 Redacted

	- Provide GP Practice information pack to provide reassurance and reduce burden regarding patient information requirements				
There is a risk that the current data controllership position as it relates to developer access specifically for necessary systems integration is misunderstood, which creates concern, direct challenge, or negative public responses	- Clear information around systems integration requirement and controls included within DPIA (section 5) - Information to be built into GP information pack for full transparency	Treat	Low	Agreed between NHS England, OpenSAFELY, and GP Profession	26/05/2023 Redacted
There is a risk that patient confidence regarding NHSE's use of data is impacted due to a lack of historic or current patient awareness regarding the service and the data it uses, which could lead to negative sentiment and increased opt-outs	- GPs provided with transparency materials and privacy notices - No planned public comms for this work, but anything further will involve public comms - Security, transparency and broad support for service should also mitigate a large proportion of this risk if patients are concerned upon initial knowledge of service existence	Treat	Medium	Agreed between NHS England, OpenSAFELY, and GP Profession	26/05/2023 Redacted
There is a risk that the data held within the system could be re-identified or used to identify patients, which could lead to a breach of GDPR and legal challenge	- Standard access model limits all access to pseudonymised data, and only a minimal set of approved researchers (under DAA or honorary contracts) have access to the Aggregated Outputs (Level 4), which are aggregate, but yet to be disclosure controlled. Risk here deemed acceptable as risk of re-	Tolerate	Low	Agreed between NHS England and OpenSAFELY	N/A

	identification is minimal and closely controlled through platform transparency and security principles - A core set of platform developers have access to the pseudonymised GP and pseudonymised NHSE data used by the platform (all under NHSE honorary contracts) as this is required for system integration and operation. All queries against this pseudonymised data are logged by EMIS and TPP; with developer queries on the level 3 cohorts logged by the OpenSAFELY SQL runner. There is a clear segregation of duties for these individuals. VPN / Account audit logs also exist to show the user and date and time of who has logged into the environment. These developers also do not have the keys for the pseudonymised data, and do not have access to any identifiable versions of the data sets held within the system - A public log of all researcher Queries and the OpenSAFELY SQL runner queries are accessibly at https://jobs.opensafely.org/; this transparent log allows the public to see how and when patient data is processed. - Only the GPSS hold the				

	encryption keys AND have access to the pseudonymised data as data processors, but their activity is fully managed through data processing agreements and current contractual frameworks covering the underlying GP Data				
There is a risk of cybersecurity threats to the service, which could be used to enable access, use, or extraction of data from the service	- TPP and EMIS are subject to all safeguards and controls covered under existing contractual within the GPIT Frameworks, which mitigate and manage cybersecurity risks for all data held within their environments - All patient derived data (for level 1, level 2, level 3, and level 4 data) associated with OpenSAFELY is stored within the TPP and EMIS secure environments; aggregated outputs that have been cleared by output-checkers can be released from the TPP and EMIS secure environments for wider circulation and publication. - Service operates to Five Safes model, and researchers must be from verified institutions and projects must have a clear sponsor and go through ethics approval - All code relating to service usage is logged and published (https://jobs.opensafely.org/)	Tolerate	Low	Agreed between NHS England and OpenSAFELY	N/A

There is a risk that research is requested and approved that would be deemed as unethical, which could undermine public trust in the service.	- All OpenSAFELY research projects are referred to HRA for ethics review (to date HRA have always accepted a review of OpenSAFELY research projects, even though the data is pseudo/anonymous) - All OpenSAFELY service evaluations/audit/research projects are required to have line manager and/or Principle Investigator approval - All OpenSAFELY service evaluation/audit projects require senior sponsor approval (usually an NHSE individual, such as a national clinical lead, band 9, director, or the sponsor might be a member of SAGE, JCVI, etc.) - All non-NHSE OpenSAFELY service evaluation/audit projects require a local ethics review committee to also verify they have reviewed the project.	Tolerate	Low	Agreed between NHS England and OpenSAFELY	N/A

21. Further Actions

- The IAO should keep the DPIA under review and ensure that it is updated if there are any changes (to the nature of the processing and/or system changes)
- A redacted version of the DPIA should be made available to the public

22. Signatories

The DPIA accurately reflects the processing and the residual risks have been approved by the Information Asset Owner:

Information Asset Owner (IAO) Signature and Date

SIRO approved Redacted 30/05/2023

FOR PRIVACY, TRANSPARENCY AND ETHICS AND OFFICE OF THE DPO USE ONLY

23. Summary of high residual risks

Risk no.	High residual risk summary
N/A	No high residual risks

Summary of DPO advice:

Data Protection Officer (DPO) **Signature and Date**

I have reviewed the DPIA and risk spreadsheet – it is very comprehensive, and I can see all previous comments and queries have been addressed. As a result, I am happy to approve this document.

Redacted 2nd June 2023

ICO consultation outcome:

Office of DPO **Signature and Date**

Formal ICO consultation not required as no high residual risks present. DPIA will be shared for visibility only.

Redacted 2nd June 2023

24. Appendices.

24.1. Appendix 1

Sender	*Content*	*Pseudonymised?*	*Mode*	*Security*	*Recipient*
A report that describes the dataset flows (name and when updated) for TPP can be found here: https://github.com/opensafely/database-notebooks/blob/master/notebooks/database-builds.ipynb Details for dataset flows with EMIS will be available shortly.					
NHSE	COVID-19 patient notification service (CPNS)	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods.	TPP and EMIS
NHSE	SUS data (APCS, ECDS, OPA)	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods.	TPP and EMIS
NHS England	HealthCare Worker status / flag (via NIMS/Pinnacle)	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard	TPP and EMIS

				methods .	
NHSE	Second generation surveillance system (SGSS) SGSS_AllTests_Negative SGSS_AllTests_Positive SGSS_Negative SGSS_Positive	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods .	TPP and EMIS
NHSE	Master Patient Index (**MPI**) with frailty flag data mart which includes pseudo **UPRN** data	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods .	TPP and EMIS
NHSE (therapeutics)	Bluteq data recording C19 antiviral / monoclonal antibody treatment for high-risk patients	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods .	TPP and EMIS
NHS England	Waiting list data **WL_ClockStops** **WL_Diagnostics**	YES	Encrypted electronic transmission, using industry standard	Encrypted file of de-identified data, encrypted using	TPP and EMIS

	WL_OpenPathways		encryption methods.	industry standard methods .	
ICBs and NHSE (via DSCROs)	High-Cost Drugs (Obtained via DARS process)	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods .	TPP and EMIS
ONS	Patient data from ONS Death records (since Feb 2020)	YES	Encrypted electronic transmission, using industry standard encryption methods. (access to ONS Portal)	Encrypted file of de-identified data, encrypted using industry standard methods .	TPP and EMIS
ONS	COVID-19 infection survey data	YES	Encrypted electronic transmission, using industry standard encryption methods. (access to ONS Portal)	Encrypted file of de-identified data, encrypted using industry standard methods .	TPP and EMIS
University of Oxford	ISARIC WHO Clinical Characterisation Protocol for Severe Emerging Infections UK (CCP-UK)	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry	TPP and EMIS

			(access to ONS Portal)	standard methods.	
UK Kidney Association's Renal Register	Renal Register	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods.	TPP
University of Leicester and University of Edinburgh	PHOSP – Post-hospitalisation COVID-19 study: a national consortium to understand and improve long-term health outcomes (PHOSP-COVID)	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods.	TPP and EMIS
King's College London*	Demographic; CSS biobank long COVID questionnaire; CSS biobank lab results (COVID-19 antibody tests); Zoe COVID symptom study app data.	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard methods.	TPP and EMIS
London School of Hygiene and Tropical Medicine (OpenPROMPT)	Patient questionnaire data demographic, symptoms, quality of life, productivity, experience of COVID	YES	Encrypted electronic transmission, using industry standard encryption methods.	Encrypted file of de-identified data, encrypted using industry standard	TPP and EMIS

				methods .	

Note: Sources marked with a "" mean that a Data Sharing agreement has been signed but data is not yet flowing*

24.2. Appendix 2

REDACTED

24.3. Appendix 3

Ongoing Approved Projects list – Covered under transitionary arrangement to retain access to Type 1 data to ensure integrity of study results

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
2	https://www.opensafely.org/approved-projects/#project-2	Online consultations	Service Evaluation	Martina Fonseca	NHS England
4	https://www.opensafely.org/approved-projects/#project-4	The effectiveness of COVID-19 vaccines against clinical outcomes: a test negative case control study using OpenSAFELY	Research	Louise Lansbury	University of Nottingham
8	https://www.opensafely.org/approved-projects/#project-8	Examining the effects of randomising dissemination of the Germ Defence website via GP practices on rates of respiratory infection, including COVID-19 and seasonal flu	Research	Jeremy Horwood	University of Bristol
9	https://www.opensafely.org/approved-projects/#project-9	What are the patterns of healthcare use in children and young people with SARS-CoV-2 in the 12 months following infection?	Research	Olivia Swann	ISARIC
11	https://www.opensafely.org/approved-projects/#project-11	Assessing the impact of COVID-19 on antidepressant prescribing	Service Evaluation	Nisha Rajendran	NHS England
12	https://www.opensafely.org/approved-projects/#project-12	Investigating events following SARS-CoV-2 infection	Research	Venexia Walker, Rochelle Knight, Jonathan Sterne	University of Bristol
16	https://www.opensafely.org/approved-projects/#project-16	How has COVID-19 impacted primary care treatment pathways for common colds, antibiotic	Service Evaluation	Victoria Palin	University of Manchester

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
		prescribing and antibiotic exposure?			
17	https://www.opensafely.org/approved-projects/#project-17	Effects of COVID-19 on people with diabetes	Research	Matthew Carr	University of Manchester
18	https://www.opensafely.org/approved-projects/#project-18	Understanding to what extent the actual treatment of people with ongoing symptoms of COVID-19 matches published NICE guidance?	Audit	Robert Willans	NICE
19	https://www.opensafely.org/approved-projects/#project-19	How has the COVID-19 pandemic impacted the need for targeted weight management interventions amongst adults in the UK? An evaluation of pandemic-associated changes in BMI and metabolic parameters.	Service Evaluation	Miriam Samuel	Queen Mary's University London
21	https://www.opensafely.org/approved-projects/#project-21	Direct Oral Anticoagulant (DOAC) prescribing during COVID-19.	Service Evaluation	Carol Roberts	PrescQIPP CIC
22	https://www.opensafely.org/approved-projects/#project-22	Investigating the effectiveness of the COVID-19 vaccination programme in the UK.	Research	Elsie Horne, Venexia Walker, Jonathan Sterne	University of Bristol
24	https://www.opensafely.org/approved-projects/#project-24	Incidence of mental illness following coronavirus infection in the community.	Service Evaluation	Daniel Ayoubkhani	Office for National Statistics
25	https://www.opensafely.org/approved-projects/#project-25	Representativeness of TPP data.	Audit	Colm Andrews	University of Oxford and LSHTM
26	https://www.opensafely.org/approved-projects/#project-26	Uptake of 'NHS @home' interventions during COVID-19.	Service Evaluation	Andy Weaver	NHS England
27	https://www.opensafely.org/approved-projects/#project-27	The effect of COVID-19 on pancreatic cancer diagnosis and care.	Service Evaluation	Agnieszka Lemanska	University of Surrey & University of Oxford
28	https://www.opensafely.org/approved-projects/#project-28	Androgen deprivation therapy (ADT) for prostate cancer and COVID-19	Service Evaluation	Agnieszka Lemanska	University of Surrey & University of Oxford
29	https://www.opensafely.org/approved-projects/#project-29	Monitoring Attendance Investigations Referrals and Outcomes related to Cancer in the NHS through the COVID-19 era (MAINROUTE).	Service Evaluation	Brian D Nicholson, Clare R Bankhead, Eva J Morris	University of Oxford
30	https://www.opensafely.org/approved-projects/#project-30	Pathology Comparators Short Report	Audit	Helen Curtis	University of Oxford and LSHTM
31	https://www.opensafely.org/approved-projects/#project-31	Risk factors and prediction models for Long COVID.	Research	Yinghui Wei	University of Plymouth & University of Bristol

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
34	https://www.opensafely.org/approved-projects/#project-34	Deaths at home during Covid-19 and implications for patients and services.	Service Evaluation		Nuffield Trust
35	https://www.opensafely.org/approved-projects/#project-35	Impact of the Covid-19 pandemic on prevention of cardiovascular disease.	Audit	Ana-Catarina Pinho-Gomes, Robert Willans	NICE
36	https://www.opensafely.org/approved-projects/#project-36	Impact of the Covid-19 pandemic on anticoagulation for atrial fibrillation.	Audit	Ana-Catarina Pinho-Gomes, Robert Willans	NICE
57	https://www.opensafely.org/approved-projects/#project-57	OpenSAFELY NHS Service Restoration Observatory: describing trends and variation in key measures of general practice clinical activity.	Research	Louis Fisher, Helen Curtis, Richard Croker, Brian Mackenna	University of Oxford and LSHTM
61	https://www.opensafely.org/approved-projects/#project-61	Estimating COVID-19 vaccine effectiveness in early recipients of the vaccine under the UK national vaccination programme	Research	William Hulme	University of Oxford and LSHTM
65	https://www.opensafely.org/approved-projects/#project-65	Describing trends and variation in HbA1c levels during COVID-19 in 24 million patients' primary care records using OpenSAFELY	Research	Robin Park	University of Oxford and LSHTM
66	https://www.opensafely.org/approved-projects/#project-66	Changes in PINCER indicators throughout the COVID-19 pandemic	Research	Louis Fisher, Lisa Hopcroft	University of Oxford and LSHTM
67	https://www.opensafely.org/approved-projects/#project-67	Assessing effectiveness of 3rd / booster COVID-19 vaccine doses in England	Research	William Hulme	University of Oxford and LSHTM
70	https://www.opensafely.org/approved-projects/#project-70	Lone households and mental health outcomes	Research	Emily Herrett	University of Oxford and LSHTM
71	https://www.opensafely.org/approved-projects/#project-71	OpenPROMPT	Research	Emily Herrett	University of Oxford and LSHTM
75	https://www.opensafely.org/approved-projects/#project-75	Post-COVID health impacts	Research	Kathryn Mansfield, Laurie Tomlinson	University of Oxford and LSHTM
77	https://www.opensafely.org/approved-projects/#project-77	Investigating the impact of COVID-19 on childhood immunisations - prototyping a dashboard	Research	Helen Curtis	University of Oxford and LSHTM
78	https://www.opensafely.org/approved-projects/#project-78	Long-term kidney outcomes after COVID-19	Research	Viyaasan Mahalingasivam, Laurie Tomlinson, Kathryn Mansfield	University of Oxford and LSHTM
83	https://www.opensafely.org/approved-projects/#project-83	Long-term kidney outcomes after SARS-CoV-2 infection	Research	Viyaasan Mahalingasivam	University of Oxford and LSHTM

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
84	https://www.opensafely.org/approved-projects/#project-84	Fit notes in primary care post-COVID	Research	Alex Walker	University of Oxford and LSHTM
85	https://www.opensafely.org/approved-projects/#project-85	FLUCATS	Research	Calum Semple	University of Liverpool
86	https://www.opensafely.org/approved-projects/#project-86	Understanding the impacts of healthcare disruption on avoidable hospitalisations	Audit	Mark Green	University of Liverpool
87	https://www.opensafely.org/approved-projects/#project-87	Validation of the OpenSAFELY kidney codes	Audit	Dorothea Nitsch, Laurie Tomlinson	University of Oxford and LSHTM
88	https://www.opensafely.org/approved-projects/#project-88	Impact of reduced kidney function, dialysis and transplantation on vaccination efficacy against COVID19	Service Evaluation	Dorothea Nitsch, Laurie Tomlinson	University of Oxford and LSHTM
91	https://www.opensafely.org/approved-projects/#project-91	Coverage, effectiveness and safety of neutralising monoclonal antibodies or antivirals for non-hospitalised patients with COVID-19	Service Evaluation	Millie Green	University of Oxford and LSHTM
92	https://www.opensafely.org/approved-projects/#project-92	The impact of the COVID-19 pandemic on services for people with asthma	Service Evaluation	Rose Higgins	University of Oxford and LSHTM
93	https://www.opensafely.org/approved-projects/#project-93	Lockdowns and vulnerable groups	Research	John Macleod, Jonathan Sterne	University of Bristol
94	https://www.opensafely.org/approved-projects/#project-94	PostOpCovid	Research	Colin Crooks	University of Nottingham
95	https://www.opensafely.org/approved-projects/#project-95	COVID-19 Collateral	Research	Rohini Mathur	University of Oxford and LSHTM
97	https://www.opensafely.org/approved-projects/#project-97	Primary Care Covid Codes Study	Audit	Colm Andrews	University of Oxford and LSHTM
99	https://www.opensafely.org/approved-projects/#project-99	Suicide deaths after healthcare contact in primary care	Audit	Hassan Ismail	NHS England
100	https://www.opensafely.org/approved-projects/#project-100	Management of Early Inflammatory Arthritis during the COVID-19 Pandemic	Service Evaluation	James Galloway	King's College London
101	https://www.opensafely.org/approved-projects/#project-101	Explaining the differential severity of COVID-19 between Indians in India and the UK	Research	Rohini Mathur	University of Oxford and LSHTM
102	https://www.opensafely.org/approved-projects/#project-102	The impact of COVID-19 on prescribing of long-term and repeated antimicrobials in primary care to evaluate antimicrobial stewardship interventions	Service Evaluation	Eleanor Harvey, Diane Ashiru-Oredope	UK Health Security Agency
103	https://www.opensafely.org/approved-projects/#project-103	Covid Symptom Study Biobank – Long Covid	Research	Nathan Cheetham	King's College London
104	https://www.opensafely.org/approved-projects/#project-104	Predictors of asymptomatic versus	Research	Thomas Cowling, Laurie Tomlinson	University of Oxford and LSHTM

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
		symptomatic infection with SARS-CoV-2			
105	https://www.opensafely.org/approved-projects/#project-105	Antidepressant trends in those with Learning Disability and Autism and changes during Covid	Service Evaluation	Christine Cunningham, Orla Macdonald	University of Oxford and LSHTM
106	https://www.opensafely.org/approved-projects/#project-106	Effectiveness and safety of sotrovimab and molnupiravir for prevention of severe COVID-19 outcomes	Research	Bang Zheng	University of Oxford and LSHTM
108	https://www.opensafely.org/approved-projects/#project-108	Hepatitis in Children related to the pandemic	Research	Brian MacKenna	University of Oxford and LSHTM
110	https://www.opensafely.org/approved-projects/#project-110	COVID-19 vaccine coverage and effectiveness in chronic kidney disease patients	Service Evaluation	Edward Parker	University of Oxford and LSHTM
112	https://www.opensafely.org/approved-projects/#project-112	The impacts of the coronavirus pandemic on rheumatoid arthritis care and non-COVID outcomes	Research	Ruth Costello	University of Oxford and LSHTM
113	https://www.opensafely.org/approved-projects/#project-113	Shared care monitoring: the impact of COVID-19 and factors associated with resilience	Research	Andrew Brown	Sunderland CCG, University of Oxford, & LSHTM
114	https://www.opensafely.org/approved-projects/#project-114	Ethnicity Short Data Report	Short Data Report	Colm Andrews	University of Oxford and LSHTM
115	https://www.opensafely.org/approved-projects/#project-115	Effectiveness of sotrovimab/molnupiravir use vs non-use	Research	John Tazare, Linda Nab	University of Oxford and LSHTM
116	https://www.opensafely.org/approved-projects/#project-116	Further analysis of DOAC prescribing for patients with mechanical heart valves	Service Evaluation	Matthew Webb	NHS England
117	https://www.opensafely.org/approved-projects/#project-117	Understanding and adjusting for bias in OpenSAFELY COVID testing data	Research	Emily Herrett, Sarah Walker	University of Oxford, LSHTM, & Office for National Statistics
118	https://www.opensafely.org/approved-projects/#project-118	Expected Background Rates of Clinical Events in Immunocompromised Patients	Research	John Tazare	University of Oxford and LSHTM
119	https://www.opensafely.org/approved-projects/#project-119	The healthcare utilisation of people with long COVID	Research	Liang-Yu Lin	University of Oxford and LSHTM
120	https://www.opensafely.org/approved-projects/#project-120	Long-term kidney outcomes following hospitalisation with Covid-19	Research	Laurie Tomlinson, Michael Sullivan	University of Oxford and LSHTM
121	https://www.opensafely.org/approved-projects/#project-121	PHOSP	Research	Max Henderson, Farag Shuweihdi	University of Leeds
122	https://www.opensafely.org/approved-projects/#project-122	Opioid prescribing trends and changes during COVID-19	Research	Andrea Schaffer	University of Oxford and LSHTM

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
125	https://www.opensafely.org/approved-projects/#project-125	What was the impact of the COVID-19 pandemic on medication reviews?	Service Evaluation	Chris Wood, Vicky Speed	University of Oxford and LSHTM
126	https://www.opensafely.org/approved-projects/#project-126	The impact of the COVID-19 pandemic on Antipsychotic Prescribing in at risk groups	Research	Orla Macdonald	Oxford Health NHS Foundation Trust
127	https://www.opensafely.org/approved-projects/#project-127	Prostate cancer incidence and prevalence in the COVID-19 pandemic	Service Evaluation	Agnieszka Lemanska	University of Surrey
128	https://www.opensafely.org/approved-projects/#project-128	The effect of COVID-19 infection and vaccination on PSA levels in men	Research	Agnieszka Lemanska	University of Surrey
129	https://www.opensafely.org/approved-projects/#project-129	Curation of GP appointments data	Audit	Lisa Hopcroft	University of Oxford
130	https://www.opensafely.org/approved-projects/#project-130	Surveillance of psychiatric emergencies over time to investigate the effect of COVID-19 on mental health. A cohort study using OpenSAFELY	Audit	Agnieszka Lemanska	University of Surrey
131	https://www.opensafely.org/approved-projects/#project-131	SLE and COVID-19	Research	Sinead Brophy	Swansea University
132	https://www.opensafely.org/approved-projects/#project-132	Assessing the Impact of COVID-19 on QOF registered patients and health groups that experience inequalities	Service Evaluation	Daniel Murdy, Daniel Shaw, Quentin Harris	NHS England
133	https://www.opensafely.org/approved-projects/#project-133	Effect of COVID-19 on prescribing of Dependence Forming Medicines and the associated health utilisation	Service Evaluation	Carol Roberts	PrescQIPP CIC
134	https://www.opensafely.org/approved-projects/#project-134	Impact of COVID-19 in people with chronic lymphocytic leukaemia	Research	Caroline Walters	University of Oxford
135	https://www.opensafely.org/approved-projects/#project-135	Impact of COVID-19 on diabetes care in England	Service Evaluation	Caroline Walters, Victoria Speed	University of Oxford
136	https://www.opensafely.org/approved-projects/#project-136	GP appointments during COVID	Service Evaluation	Lisa Hopcroft	University of Oxford
137	https://www.opensafely.org/approved-projects/#project-137	Healthcare needs for people with chronic kidney disease in the COVID-19 era	Research	Viyaasan Mahalingasivam, Laurie Tomlinson	London School of Hygiene and Tropical Medicine
138	https://www.opensafely.org/approved-projects/#project-138	Hypertension and blood pressure in the Quality and Outcomes Framework (QOF)	Service Evaluation	Milan Wiedemann	University of Oxford
139	https://www.opensafely.org/approved-projects/#project-139	Risk factors for COVID-19 disease progression in immunocompromised populations	Research	Edward Parker	London School of Hygiene and Tropical Medicine
140	https://www.opensafely.org/approved-projects/#project-140	Incidence and management of gout	Service Evaluation	James Galloway	King's College London

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
		during and before the COVID-19 pandemic			
141	https://www.opensafely.org/approved-projects/#project-141	The impact of COVID-19 on the care of people with sickle cell disease.	Service Evaluation	Brian MacKenna	University of Oxford
142	https://www.opensafely.org/approved-projects/#project-142	The impact of COVID-19 on prescribing of antimicrobials	Service Evaluation	Brian MacKenna	University of Oxford
143	https://www.opensafely.org/approved-projects/#project-143	Comparison of risk factors for hospitalizations and death from winter infections	Research	Venexia Walker, Rachel Denholm, Jonathan Sterne, Angela Wood, Ben Goldacre	University of Bristol, University of Cambridge, & University of Oxford
144	https://www.opensafely.org/approved-projects/#project-144	COVID-19 impact on anastrazole use by post-menopausal women with hormone-dependent breast cancer	Service Evaluation	Darren Williams	NHS England
145	https://www.opensafely.org/approved-projects/#project-145	Association between ursodeoxycholic acid and COVID-19 outcomes	Research	Ruth Costello	London School of Hygiene and Tropical Medicine
146	https://www.opensafely.org/approved-projects/#project-146	Scarlet fever and invasive group A strep cases during the COVID-19 pandemic	Service Evaluation	Brian MacKenna	University of Oxford
147	https://www.opensafely.org/approved-projects/#project-147	Winter pressures in primary care in the context of COVID-19 recovery	Service Evaluation	Alex Walker	University of Oxford
148	https://www.opensafely.org/approved-projects/#project-148	The impact of COVID-19 on pregnancy treatment pathways and outcomes	Research	Victoria Palin	University of Manchester
149	https://www.opensafely.org/approved-projects/#project-149	Descriptive cohort analysis of Long COVID and vaccination status	Research	Alasdair Henderson	London School of Hygiene and Tropical Medicine
150	https://www.opensafely.org/approved-projects/#project-150	Evaluating the UK ‘shielding’ policy during the COVID-19 pandemic	Research	Johnny Filipe	London School of Hygiene and Tropical Medicine
151	https://www.opensafely.org/approved-projects/#project-151	Venous thromboembolism and cardiovascular events in inflammatory rheumatic diseases	Research	James Galloway	King's College London
152	https://www.opensafely.org/approved-projects/#project-152	Digital access to primary care for older people during COVID	Research	Abodunrin Aminu, David Sinclair, Alex Hall, Andrew Clegg, Barbara Hanratty, Chris Todd	University of Manchester & Newcastle University
153	https://www.opensafely.org/approved-projects/#project-153	What was the impact of COVID-19 on mortality related to venous	Service Evaluation	Dr Matthew Beresford	NHS England

Project Number	Link to project on approved projects page	Name/Title	Research Type	Research Owner	Organisation
		thromboembolism in England?			
154	https://www.opensafely.org/approved-projects/#project-154	Validation of ISARIC / SUS / PHOSP data for COVID-related hospital admissions	Research	Will Hulme, Alex Walker	University of Oxford
155	https://www.opensafely.org/approved-projects/#project-155	The effect of Long COVID on Quality Adjusted Life Years using OpenPROMPT	Research	Andrew Briggs, Oliver Carlile	London School of Hygiene and Tropical Medicine