Data Provision Notice

The NHS OpenSAFELY Data Analytics Service Pilot

Information Asset Owner: Director of Data Access and Partnerships

Version: 1.0

Publication date: July 2025

Copyright © 2025 NHS England

Contents

Background 3

Purpose of the Pilot 3

Purpose of this Data Provision Notice 4

Benefits 4

Organisations consulted. 5

Legal basis for the collection 6

Organisations within scope of this Data Provision Notice 6

Scope of the Collection 6

Required Data and the Form, Manner and Timing of the Data to be Collected 8

Period of the Collection 8

Burden of the collection 8

Steps taken by NHS England to minimise the burden of collection 8

Copyright © 2025 NHS England

Background

The Health and Social Care Act 2012 (the **2012 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, where it has been directed to establish an information system by the Secretary of State for Health and Social Care.

Data is required, as specified by NHS England in this published Data Provision Notice below, to enable NHS England to comply with the NHS OpenSAFELY Data Analytics Service Pilot Directions 2025 (**OpenSAFELY Pilot Directions**) and Directions Requirements Specification (**Requirements Specification)** issued by the Secretary of State for Health and Social Care to NHS England. Therefore, organisations that are in scope of the notice are legally required, under section 259(5) of the Act, to provide the required data in the form and manner specified below.

The OpenSAFELY COVID-19 Service was established under the COVID-19 Public Health Directions 2020[1] (**COVID-19 Public Health Directions**) in June 2023 and provides a secure analytics service for users, approved by or on behalf of NHS England, to run queries on pseudonymised patient data held by GP practices and NHS England (**GP and NHS England pseudonymised patient data**) for COVID-19 Purposes[2].

Learning from the OpenSAFELY COVID-19 Service (the **COVID-19 Service**), it has been decided that an additional service should be developed and tested on a pilot basis to enable GP and NHS England pseudonymised patient data to be accessed through the OpenSAFELY platform for further secondary uses purposes (**Pilot**). This new Pilot will operate as described below.

For avoidance of doubt, the COVID-19 Service will continue to operate independently of this Pilot and is out of scope of this Data Provision Notice.

Purpose of the Pilot

The purpose of the Pilot is to establish a secure analytics service using the OpenSAFELY platform (**the Service**), for users approved by or on behalf of NHS England (**Approved Users**), to run queries on GP and NHS England pseudonymised patient data for the purposes of the Pilot as set out below. The list of NHS England pseudonymised data sets held within the Service is set out in **Appendix 1**.

The Service, and the purpose of running the queries and obtaining the aggregate outputs is limited to the following purposes:

- clinical audit[3]
- service evaluation[4]
- health surveillance[5]

[1] COVID-19 Public Health Directions 2020 - NHS England Digital
[2] See paragraph 2.2 of the Covid-19 Public Health Directions 2020 as to the Covid-19 Purposes.
[3] As defined in the https://www.hra-decisiontools.org.uk/research/docs/DefiningResearchTable_Oct2022.pdf
[4] As defined in the https://www.hra-decisiontools.org.uk/research/docs/DefiningResearchTable_Oct2022.pdf
[5] As defined in the https://www.hra-decisiontools.org.uk/research/docs/DefiningResearchTable_Oct2022.pdf

Copyright © 2025 NHS England

- research[6]
- health and social care policy, planning and commissioning purposes and public health purposes, where agreed on a project specific basis by or on behalf of:
    - the Deputy Director of Data Policy on behalf of the Department of Health and Social Care, and
    - the Chief Data & Analytics Officer on behalf of NHS England, and
    - a nominated representative of each of the Royal College of General Practitioners and the British Medical Association on behalf of the Joint GP IT Committee.
- evaluation of the Service.

The Service is designed to keep patient data confidential: Approved 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 data sets specifically tailored to the needs of the project (**Intermediate Outputs**). Further Queries are run to generate aggregated outputs (**the Aggregated Outputs**) and logs to help identify and fix errors in the analysis code (**Error Logs**).

Approved Users only have access to the Aggregated Outputs and Error Logs generated by their Query. Aggregated Outputs are released outside the GP System Suppliers' secure environment only after disclosure controls have been applied within the GP System Suppliers' secure environment and the results reviewed and cleared by trained output-checkers. All actions in the Service are logged in public, in real-time and all Queries are logged and published.

Purpose of this Data Provision Notice

This Data Provision Notice is being issued to GP Practices to require GP Practices to provide the Service with access to pseudonymised GP patient data:

- for system integration activities, and
- in order to run Queries against that data to produce the Intermediate and Aggregated Outputs and the Error Logs for the Purposes of the Pilot under the Directions.

NHS England, as the operator of the Service will provide Approved Users who developed the Query with secure access to the Logs and Aggregated Outputs.

Further information about the Pilot can be found in the Requirements Specification and other information published here NHS England transparency notice.

Benefits

NHS England and Approved Users need insight from this vital data to analyse and link data for research, clinical service evaluation, clinical audit and health surveillance which will help improve the quality of people's lives and help save lives. Benefits which may be realised include:

[6] As defined in paragraph 3.1 of the UK Policy Framework for Health and Social Care Research

Copyright © 2025 NHS England

- understanding risks to public health, trends in diseases and such risks, and controlling and preventing the spread of communicable diseases and such risks
- identifying and understanding information about patients or potential patients with, or at risk of disease
- understanding information about patient access to health services as a direct or indirect result of illness, and the availability and capacity of those services
- health services research on changes in clinical activity and population health and impact, as well as consequences of this
- informing the development and tracking the implementation of new diagnostics and treatments.

Supporting the use of GP and NHS England data for the Purposes above via the Pilot will reduce the burden on General Practice, and the wider NHS (including hospital, community and mental health services), as well as provide increasing support to social care services; all at a time when demand on resources is high. It is hoped the insights will support General Practice and the wider health and care system to improve the quality and safety of care and support to patients.

Organisations consulted.

NHS England has, as required under section 258 of the 2012 Act, consulted with the following organisations:

- GP Representatives including from:
    - The British Medical Association
    - The Royal College of Practitioners (RCGP)
- NHS England's Advisory Group for Data (AGD)
- Citizen Juries
- UseMyData
- The National Data Guardian for Health and Social Care
- The Phoenix Partnership Ltd (Leeds) (TPP)
- Egton Medical Information Service Ltd (EMIS)
- NHS England's OpenSAFELY Oversight Board
    - Organisational Members
        - The Bennett Institute
        - Wellcome Trust
        - Open Data Institute
        - LSHTM
        - NHS England
        - RCGP
        - BMA
        - MedConfidential

Copyright © 2025 NHS England

- EMIS
- TPP
- Association of Professional Healthcare Analysts (AphA CIC)

Legal basis for the collection

NHS England has been directed by the Secretary of State for Health and Social Care under the OpenSAFELY Pilot Directions to establish and operate an information system for this Service. The OpenSAFELY Pilot Directions and accompanying Requirements Specification are published on the NHS England website: https://digital.nhs.uk//about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/secretary-of-state-directions/nhs-opensafely-data-analytics-service-pilot-directions-2025

Organisations within scope of this Data Provision Notice

Under section 259(1)(a) of the 2012 Act, this Notice is served in accordance with the procedure published as part of the NHS England duty under section 259(8) of the 2012 Act **on all General Practices across England using TPP and Optum (formerly EMIS Group PLC) health record systems**.

Under section 259(5) of the 2012 Act, General Practices within this scope must provide the data identified below as the Required Data, in the form and manner and in accordance with the timing requirements specified below for the period of the collection set out below.

Scope of the Collection

Patients within Scope

The GP patient population made available to the Service for analysis is defined as all patients EXCEPT:

- Patients who have no period of registration (in a TPP or **Optum (formerly EMIS Group PLC)** practice) after 1 January 2009 or who died before 1 January 2009 (the year of the Swine Flu Pandemic). This means that for each GP 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 who have registered a Type 1 Opt-out with their GP Practice. Notwithstanding that no identifiable patient data leaves the GP System Suppliers' secure analytics environment through the operation of the Service, Type 1 Opt-outs will be respected by the Service. This is consistent with other NHS England GP data collections directed under section 254 of the 2012 Act, for planning and research purposes.

Copyright © 2025 NHS England

In certain limited circumstances and where ethics approvals or advice from AGD following consultation with nominated representatives of BMA and RCGP support it, a project may wish to apply the National Data Opt Out (NDOO) as part of the Query they have developed, notwithstanding that the Service operates under an exemption to the National Data Opt Out Policy. NHS England will therefore work with GP System Suppliers to develop a technical solution to enable the application of the NDOO at project level at the request of an Approved User in these circumstances.

Type of Data within Scope

The full pseudonymised GP data processed by the Service to deliver Error Logs and Intermediate and Aggregated Outputs contains structured and coded data only. It does not contain free text or information such as letters, images, and correspondence. It includes:

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

For this patient population within scope above, all pseudonymised GP patient data will be made available for analysis, EXCEPT the SNOMED CT codes in reference sets reflecting current guidance regarding restrictions around the sharing of certain patient information that is legally restricted and sensitive.[7]

Categories of Data within Scope

The categories of personal data and special category data which will be made available to the Service are set out in the Requirements Specification.

Source of Data within Scope

The data sources for the Service are:

1. **General Practice Datasets** - Pseudonymised GP patient records (only the coded information - with numerical values where applicable) stored in TPP or EMIS' systems as detailed above.
2. **NHS England Datasets** - NHS England pseudonymised patient datasets which NHS England holds as Controller and which are listed in **Appendix 1**. NHS England may add other datasets to the Service where it considers it necessary to achieve the Purpose of the Pilot and this is approved in writing by or on behalf of NHS England's Chief Data & Analytics Officer and Senior Information Risk Officer.

[7] ● 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'.

The above-restricted codes reflect current guidance regarding restrictions around the sharing of certain patient information and can be updated to align with new NHS England and professional guidance.

Copyright © 2025 NHS England

More information about the datasets used as part of the Service is available on the OpenSAFELY website here https://docs.opensafely.org/data-sources/

Required Data and the Form, Manner and Timing of the Data to be Collected

Required Data

The pseudonymised GP patient data alongside the pseudonymised NHS England patient data from each Query, which forms part of the Intermediate Output and Logs, is processed for the purposes of running a Query and is the data required to be collected by NHS England under the OpenSAFELY Directions and which GP Practices are required to provide under section 259(5) of the 2012 Act (the **Required Data**).

Form, Manner and Timing for collection of the Required Data

The Required Data will be provided in response to every Query that is run on the pseudonymised GP patient data at the time the Query is run.

The Intermediate Output and Logs remain within the GP System Supplier's secure analytics environment and are not released from that environment. Only the Aggregated Output will be released from the GP System Supplier's secure analytics environment as part of the Service, to the Approved User who ran the Query, and only after disclosure control checks have been applied by trained output-checkers.

Period of the Collection

The Service will be delivered under the OpenSAFELY Directions until the 31st of March 2027

Burden of the collection

Steps taken by NHS England to minimise the burden of collection

NHS England has a statutory duty under section 253(2) of **the 2012 Act** to seek to minimise the burden it imposes on others. In seeking to meet these obligations in relation to this Service, the Service has been designed to reduce the burden, responsibility, and risks on the General Practice of individually administering multiple requests for access to data, through NHS England operating this Service in collaboration with the Bennett Institute (University of Oxford).

As this is an existing Service (for COVID-19 purposes) which will now also deliver the Pilot, there is no additional burden on General Practices.

Copyright © 2025 NHS England

Appendix 1: NHS England Datasets within the Service

Number	Dataset
001	NHS England Secondary Use Service (SUS)
002	NHS Civil Registration Data in England & Wales

Copyright © 2025 NHS England