Data Provision Notice

General Practice Data for Planning and Research

Information Asset Owner: Dave Roberts

Version: 1.0

Published: 12th May 2021



Information and technology
for better health and care

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Contents

Background 4

GP Data for Planning and Research data collection 4

Purpose of the collection 5

Benefits of the collection 6

Legal basis for the collection, analysis, publication and dissemination 6

Collection and Analysis 6

Opt-outs 6

Publication 8

Dissemination 8

Processing of personal data 9

Processing of special categories of personal data 10

Transparency 11

Health and Social Care Bodies within the scope of the collection 12

Form and manner of the collection 12

Form of the collection 12

Manner of the collection 13

Data linkage and pseudonymisation 14

Period of the collection 15

Burden of the collection 15

Steps taken by NHS Digital to minimise the burden of collection 15

Detailed burden assessment findings 16

Persons consulted 16

Appendix A – Data Minimisation, Additional Protections and Pseudonymisation 18

Data minimisation and additional protections 18

Data for Direct Care (by exception) 19

Pseudonymisation 19

Pseudonymisation tool 21

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Appendix B – Data Model 22

Appendix C – NHS Digital dissemination of General Practice Data 23

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Background

The General Practice Extraction Service (GPES) has now been running for over 10 years. It currently undertakes over 350 extracts each year, the majority of which are concerned with payments to General Practice. The COVID-19 pandemic has also put additional pressures on the service. The technology used by the service is old and there is an enormous amount of human resource involved from both NHS Digital and GP system suppliers which makes the service unsustainable going forward.

In 2014, a National Audit Office (NAO) review of GPES found that the service was inefficient and costly but fulfilled a necessary requirement and should be replaced and improved. A subsequent Public Accounts Committee (PAC)[1] hearing in Parliament upheld and supported these findings. Whilst it is legally compliant, and has worked well for over 10 years, GPES is inefficient, costly, and capacity constrained because it requires a new collection for each new requirement, a “collect once, use once” method, with limited automation.

The Health and Social Care Act 2012 (the 2012 Act) gives the Health and Social Care Information Centre, now known as NHS Digital[2] and hereafter referred to by this name, 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 (Secretary of State) or NHS England and requires that data to comply with a direction.

GP Data for Planning and Research data collection

In order to provide for the replacement of GPES, the Secretary of State has directed NHS Digital to establish an information system to collect and analyse General Practice data to support the purposes set out in the GP Data for Planning and Research Direction 2021 (the Directions) and noted below. This will enable NHS Digital to provide access to the General Practice data, through the NHS Digital Data Access Request Service (DARS) in accordance with the processes set out in more detail below, to organisations who have a clear legal basis and need to access it for purposes outlined in the Directions. It will also enable NHS Digital to publish relevant statistical data where either this has been agreed with the Department of Health and Social Care first, or where the Chief Statistician of NHS Digital considers this would be in accordance with the Code of Practice for Statistics.

In addition to a phased retirement of GPES, this new information system is required to:

- provide capacity to meet the rapidly increasing requirements for different types of data from patient level identifiable to publishable statistics to support local, regional, and national planning, policy development, public health (including COVID-19), commissioning and research.

[1] https://old.parliament.uk/business/committees/committees-a-z/commons-select/public-accounts-committee/inquiries/parliament-2015/general-practice-extraction-service-2015-16
[2] https://digital/nhs.uk/

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

- achieved in a way that builds and maintains professional, patient, and public trust, through adherence to UK GDPR [3] (GDPR), the Data Protection Act 2018 (DPA) and the Human Rights Act 1998 and operates within the legal, policy and transparency framework of NHS Digital, including its functions and powers to share data set out in the 2012 Act and compliance with the common law duty of confidentiality.

The data, as specified by NHS Digital in this published Data Provision Notice (DPN or Notice), is required to support this Direction. Therefore, organisations that are in scope of the Notice are legally required, under sections 259(1)(a) and 259(5) of the 2012 Act, to provide the data as specified in the form and manner sections below.

NHS Digital has engaged with representatives of the British Medical Association (BMA), the Royal College of General Practitioners (RCGP) and the National Data Guardian (NDG) to design an improved approach for the collection and controlled access to data from General Practice.

This improved approach involves the collection of patient data which is pseudonymised by GP system suppliers at source before it is shared with NHS Digital. It also meets the ever-increasing demand for access to GP data by trusted organisations in a safe and secure way. This approach will help to support local, regional, and national planning, policy development, public health, commissioning and research designed to improve health and care treatment and services for patients and will relieve the burden on General Practice by reducing the number of requests for data extractions, form filling and submissions being requested of GP Practices.

The transition to the new system to enable the replacement of GPES will take place over a period of 18 months.

This Notice is issued in accordance with the procedure published as part of NHS Digital's duty under section 259(8) of the 2012 Act.

Purpose of the collection

The purpose of the data collection is to support the provision of General Practice data for health and social care purposes including supporting local, regional, and national planning, policy development, public health, commissioning, and research.

The data collection is designed for the above secondary uses, however certain direct care purposes may require this data to contribute to a clinical intervention. Further details relating to this can be found in the Appendix A Data for Direct Care (by exception).

NHS Digital will undertake a managed transition from GPES to the new GP data service. NHS Digital will also work with other organisations to enable transition from existing data flows to the new NHS Digital GP Data Service in order to reduce the number of GP data flows currently in existence. This transition needs to be carefully managed, so in the meantime, General Practices should continue to support existing lawful data flows. New requests for access to data from 1st September 2021 for planning and research should be directed to NHS Digital.

[3] As defined in section 3(10) of the Data Protection Act 2018

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Benefits of the collection

The pseudonymised data collection follows the 'extract once, utilise many times' approach, where data from a single collection can, through the NHS Digital Data Access Request Service, be used for multiple approved and lawful purposes to improve health and care services. This will:

- reduce the burden, responsibility, and risks on General Practice of individually administering multiple requests for access to data
- simplify the route for third parties to seek access to data from General Practice and reduce the burden on resources of GP system suppliers and NHS Digital
- apply consistent protections in transit and when stored by NHS Digital for personal data from GP medical records through using pseudonymised data by default
- simplify and improve the transparency of information to patients about how data from General Practice is used to help run and improve health and care services in England through providing patients with information in the GP Privacy Notice, the NHS Digital Transparency Notice, the NHS Digital Data Release Register and case studies on NHS Digital's website about how data has been used.
- establish a service that can meet the increasing demand for access to data from General Practice in an efficient, safe, consistent, timely and transparent way

Legal basis for the collection, analysis, publication and dissemination

Collection and Analysis

NHS Digital has been directed by the Secretary of State under section 254 of the 2012 Act under the General Practice Data for Planning and Research Directions 2021 to establish and operate a system for the collection and analysis of the information specified within the Directions. A copy of the Directions is published here: General Practice Data for Planning and Research Directions 2021[4] (the Directions).

Appendix B – Data Model of this Notice provides a link to a diagram which shows a high-level view of the data items being collected and how these can be linked across the various data tables. The relevant data items that directly identify the patient will be pseudonymised before leaving the GP systems in order to protect the security of patients' data. Data will only be re-identified for approved specific uses where pseudonymised data would not be adequate for the purpose and where the law allows. There is more about this in the Dissemination section below.

Opt-outs

Patients may have registered a Type 1 opt-out or a national data opt-out choice concerning the use of their identifiable data for purposes beyond their individual care.

[4] https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/secretary-of-state-directions/general-practice-data-for-planning-and-research-directions-2021

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

A Type 1 opt-out (also known as a Type 1 objection) prevents an individual's identifiable patient data from being shared outside of their General Practice except when it is being used for the purposes of their individual care. In line with the current policy on Type 1 opt-outs set out in the Government response to the National Data Guardian (NDG) review of data security, consent and opt-outs, NHS Digital will uphold Type 1 opt-outs. NHS Digital will therefore not collect data for any patients who have a Type 1 opt-out registered with their GP Practice, from the date that Type 1 opt-out is registered. This may change in the future if NHS Digital is directed otherwise in the event of a change in policy following a review of Type 1 opt-outs by NHSx, with implementation being subject to consultation with the profession via the Joint GP IT Committee, a representative body comprised of elected members from RCGP and BMA.

It may also change if NHS Digital agrees with the British Medical Association (BMA) and the Royal College of General Practitioners (RCGP), and the Department of Health and Social Care (DHSC) that it has put in place appropriate organisational and technical measures and controls to enable it to collect and process pseudonymised Type 1 opt-out records by means which continue to uphold the Type 1 opt-out and do not enable the patient to be directly identified except for the purposes of their own care.

GP system suppliers will commence processing of data and extractions for individual GP Practices who have responded to their system supplier to confirm they are complying with this Notice seven weeks from the date of issue of this Notice. This means no processing and no data will be extracted by GP system suppliers and provided to NHS Digital before the 1st July 2021.

The standard notice period provided for a GPES collection is six weeks from the point the DPN is issued. The GP Privacy Notice and the NHS Digital Transparency Notice reflect this six week window, so that if a patient wishes to register a Type 1 opt-out to prevent their information being shared, that they can do so by registering the Type 1 opt-out with their GP Practice by downloading a form from NHS Digital here.

This provides GP Practices with a week to apply any Type 1 Opt-out requested by a patient, before 1st July 2021 when data extractions will commence, to ensure there is time for the Type 1 opt-out to be honoured, and that no data is shared by the GP Practice with NHS Digital for those patients.

If patient data has already flowed to NHS Digital before a Type 1 opt-out is registered, the data already held by NHS Digital will continue to be accessible. This is because NHS Digital cannot identify the records of those who have registered a Type 1 opt-out. However, no new data will be collected about a patient following the registration of a Type 1 opt-out. Patients can also register a National Data Opt-out which will be applied by NHS Digital in accordance with the National Data Opt-out policy in relation to any requests to grant access to the data that is already held.

Patients can download a form NHS Digital has produced and use this to register a Type 1 opt-out with their GP Practice. A link to this form is in the GP Privacy Notice and the NHS Digital Transparency Notice on the NHS Digital website. More information is set out in the Transparency section below.

The National Data Opt-Out does not apply to any collection of data by NHS Digital, however, it will be applied by NHS Digital on access or dissemination of data, in accordance with the National Data Opt-out Operational Policy Guidance.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Publication

NHS Digital may publish statistical data which is anonymous derived from this collection where this is agreed with the Department for Health and Social Care or where the Chief Statistician of NHS Digital considers a publication is appropriate. These publications will be in accordance with the Code of Practice for Statistics. Any publication of this information would be discussed with the British Medical Association and Royal College of General Practitioners.

Any information that is published will be fully anonymised for publication in accordance with the Information Commissioner's Office Anonymisation Code of Practice[5] and any successor guidance issued by the ICO and be in accordance with the Code of Practice for Statistics.

Dissemination

NHS Digital has responsibility and accountability at all times for the dissemination of data from the collection as the Data Controller under the GDPR. Dissemination is the legal term given to NHS Digital's functions to share data under the 2012 Act and covers providing access to data within NHS Digital's secure environment and sharing data by way of an extract of data.

NHS Digital will ensure that approved requests for access to the data are necessary, proportionate, that the minimum amount of data necessary for the purpose only is shared and that the access and use of the data shared will be secure and lawful.

NHS Digital will hold the collected data securely and only provide access to it following robust review and approval through the Data Access Request Service (DARS). This could be as pseudonymised patient level or aggregated data, or identifiable patient level data where legally permitted and necessary for the purpose, such as for clinical trials. All requests for data from this service will be managed through DARS. Exceptions are requests for anonymous, aggregate statistical data which will be dealt with by the NHS Digital Chief Statistician and requests for access for direct care, which will be approved through a separate NHS Digital direct care approval process, which includes clinical assurance and approval by the Caldicott Guardian. The DARS process is robust and well established, and consists of enquiry, triage, review, independent oversight through the Independent Group Advising on the Release of Data (IGARD), approval by the NHS Digital Senior Information Risk Owner (SIRO), including access, audit and destruction phases.

Data released through DARS will, where possible, be provided to data recipients through a secure data access environment (DAE) within NHS Digital infrastructure. Where the needs of the recipient cannot be met this way, and the recipient meets the DARS security standards, there will be a direct dissemination of a data extract. NHS Digital plan to reduce the amount of data being processed outside central, secure data environments and increase the data we make available to be accessed via our secure data access environment. All data approved for release through DARS and IGARD are subject to robust data sharing agreements between NHS Digital and the organisation requesting the data. The data sharing agreements include provisions which enable NHS Digital to audit data usage and to terminate the data sharing agreement, revoke data access and require destruction of the data where there are breaches

[5] https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf or any subsequent document on the same topic published by the ICO. Note this Code has not been withdrawn by the ICO and although it is not up to date and reflective of UK GDPR, it still remains relevant guidance in relation to anonymisation until replaced by the ICO with updated guidance.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

of the data sharing agreement. More information about the protections in data sharing agreements is in Appendix C and about audits on our Data Sharing Audits webpage.

Requests by organisations to access record level data and any identifiable data from this collection will be subject to independent oversight by IGARD. In addition, NHS Digital will also seek the views of representatives of the BMA and the RCGP through the Professional Advisory Group as part of the DARS process.

Where identifiable patient level data is requested, data applicants will need to demonstrate they have a lawful basis to access the data for the purposes set out in the Directions and to process this data without breaching the common law duty of confidentiality. This may include express patient consent or an approval under section 251 of the National Health Service Act 2006 following support from the Confidentiality Advisory Group, for example in the case of certain research and clinical trials. Use of data for research purposes will also require a Research Ethics Committee approval.

NHS Digital discloses in its Data Release Register the organisations to whom it allows access or to which it disseminates the data obtained through this Notice, also the form and description of the data, the legal basis and the purposes of the access or dissemination. The Data Release Register also sets out whether the National Data Opt-out has been applied. NHS Digital will also publish examples of how the data from this collection has been used, to inform the public and the profession of the benefits from the use of the data.

Further details of the NHS Digital process to ensure the lawful and ethical dissemination of data is provided in Appendix C.

Processing of personal data

The data to be collected, whilst pseudonymised, is still personal data under GDPR as NHS Digital has the means to re-identify the data in accordance with the agreed processes for this. This includes review and recommendation by IGARD and the Professional Advisory Group, and a signed data sharing agreement being in place before access to identifiable data can be provided to requesting organisations.

Processing of the data collected by NHS Digital is lawful under Article 6(1)(c) of GDPR in order to comply with the Directions from the Secretary of State for Health and Social Care.

The sharing of their patients' data with NHS Digital by General Practices is also lawful under Article 6(1)(c) as General Practices have a legal obligation, under sections 259(1)(a) and 259(5) of the 2012 Act, to provide the data to NHS Digital as specified in the form and manner below.

Extract:

> ***Article 6 (1), GDPR "Lawfulness of processing"***
>
> *Processing shall be lawful only if and to the extent that at least one of the following applies:*
>
> *(c) processing is necessary for compliance with a legal obligation to which the controller is subject;*

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Processing of special categories of personal data

As the data to be collected contains data about health, which is a special category of personal data under GDPR, processing is lawful under GDPR:

(i) in the case of NHS Digital's collection of the data, under Article 9(2)(g) (substantial public interest) of GPDR and Schedule 1, Part 2, Paragraph 6 (2)(a) (statutory functions) of the Data Protection Act (DPA) 2018. This is because it is substantially in the public interest to collect and analyse the data for the purposes set out in the Direction on the basis of section 254 of the 2012 Act. It is also necessary for NHS Digital to process personal data in the exercise of the statutory function conferred on it under the Direction and s254 of the 2012 Act.

(ii) in the case of sharing of the data by General Practices, under Article 9(2)(g) (substantial public interest), Article 9(2)(h) (healthcare purposes), (i) (public health purposes) and (j) (research and statistical purposes) of GDPR and Schedule 1, Part 1, Paragraphs 2(2)(f), 3, 4 and 6(2)(a) of the DPA 2018.

Extracts from GDPR and the DPA for these legal bases are set out below:

Article 9 (2)(g), GDPR – Substantial Public Interest

Processing is necessary for reasons of substantial public interest, on the basis of domestic law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject;

Schedule 1, Part 2, Paragraph 6 (2)(a), DPA 2018 – Statutory Functions

Processing is necessary for the exercise of a function conferred on a person by an enactment or rule of law.

Article 9 (2)(h), GDPR – Healthcare

Processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of domestic law;

Schedule 1, Part 1, Paragraph 2 (2)(f), DPA 2018 - Healthcare

Processing is necessary for the management of health care systems or services or social care systems or services.

Article 9(2)(i) – Public Health

Processing is necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

standards of quality and safety of health care and of medicinal products or medical devices, on the basis of domestic law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy;

Schedule 1, Part 1, Paragraph 3, DPA 2018 – Public Health

Processing—

(a) is necessary for reasons of public interest in the area of public health, and

(b) is carried out—

(i) by or under the responsibility of a health professional, or

(ii) by another person who in the circumstances owes a duty of confidentiality

under an enactment or rule of law.

Article 9(2)(j), GDPR – Research and Statistical Purposes

Processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) (as supplemented by section 19 of the 2018 Act) based on domestic law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

Schedule 1, Part 1, Paragraph 4, DPA 2018 – Research and Statistical Purposes

Processing is necessary for scientific research purposes or statistical purposes.

Transparency

As data controllers NHS Digital and General Practices have a legal duty to be lawful, fair and transparent and to provide patients with accessible information under GDPR about the data they are sharing.

General Practices will need to update their own Privacy information before this collection commences. NHS Digital has produced an additional General Practice Privacy Notice for this collection on the NHS Digital website, which GPs can easily link to for this purpose. This Privacy Notice provides important information to patients about their rights to opt out, including their right to exercise a Type 1 Opt-out to stop their identifiable data being shared with NHS Digital for purposes beyond their direct care. Although NHS Digital is collecting data that is pseudonymised, it has agreed with the BMA and the RCGP it will not collect data about patients who have registered a Type 1 Opt-out. The Privacy Notice provides a link to a Type 1 Opt-out Form which patients can complete and send to their GP Practice by post or email. When this is received, GP Practices must action this promptly by registering the Type 1 Opt-out on the patient's record using the codes set out in the NHS Digital form **by no later**

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

than the 30th June 2021. A copy of the Form, which is hosted on NHS Digital's website is accessible here. Patients may however exercise this right by using other forms available, or by simply asking at the GP Practice.

NHS Digital has published its own NHS Digital Transparency Notice for this collection which provides more information about how NHS Digital will process the data collected from General Practice and also sets out all of the legal bases under GDPR which apply to NHS Digital's subsequent sharing of any GP data with other organisations. This also contains full information about opting out and provides a link to the Type 1 Opt-out form. The General Practice Privacy Notice contains a link to this NHS Digital Transparency Notice so that if patients would like more information, they can click through to the NHS Digital Notice.

Health and Social Care Bodies within the scope of the collection

Under section 259(1)(a) of the Health and Social Care Act 2012, this Notice is served in accordance with the procedure published as part of the NHS Digital duty under section 259(8) on the following persons:

- General Practices in England

Under section 259(1)(a) and (5) of the Health and Social Care Act 2012 General Practices must comply with the Form, Manner and Period requirements below.

Form and manner of the collection

Form of the collection

The data collection will include coded and structured data for all patients (adults and children) registered with a General Practice and controlled by GP Practice Data Controllers relating to the delivery of General Practice care at the start of the collection, including medical record and appointment administration as specified in Annex B of the General Practice Data for Planning and Research Directions 2021.

GP system suppliers will remove out of scope items such as name and address and pseudonymise the identifiable data at source before it leaves the control of their GP Practice.

Data of those patients deceased after their GP Practice has commenced flowing this data collection will be retained by NHS Digital for the duration of the Direction and in accordance with NHS Digital's Records Management Policy.

The breadth and depth of coded and structured data to be collected has been tested against current and immediate planned use cases (separate to the COVID-19 pandemic requirements). This includes coded data about physical, mental and sexual health, including terms included in legacy reference sets of 'sensitive codes' primarily relating to sexually transmitted infections[6]. This data is requested via NHS Digital or alternative channels and includes codes currently

[6] The legacy code list is still published as a technical product at this link Releases · TRUD (digital.nhs.uk) Please note this requires registration and may not meet the accessibility needs of all users.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

recorded in records and historic coded data, including for example within the following groups in the GP medical record systems.

(more information is provided in Appendix B):

- Patient demographics
- Diagnoses and symptoms
- Observations and encounters
- Test results
- Staff details
- Medications, allergies and immunisations
- Referrals and recalls
- Appointments (including appointment management and administration data)

If you want to know more about what codes are collected look here.

Items not collected are as follows:

- All coded record elements identified as being legally restricted under Fertility, Embryology and Gender recognition legislation
- All coded and structured data for patients who have a Type 1 opt-out active at their GP Practice
- All structured and coded data relating to appointments, medications and referrals more than 10 years old
- All 'free text' medical record content
- All attached documents, images and other file types
- All paper records including scanned
- All audit trail data

In the case of GP medical record systems where the architecture enables access to and viewing of data entered outside of General Practice, GP system suppliers must only extract and flow data for which GPs are the data controller either solely or jointly.

More detail is provided in Appendix A – Data Minimisation, Additional Protections and Pseudonymisation and in Appendix B – Data Model.

Manner of the collection

Invitation to comply with this Notice will NOT be via the Calculating Quality and Reporting Service (CQRS), the route usually used for GPES collections, because this collection is not being done as a part of GPES.

GP Practices will be sent an invitation to comply with the Data Provision Notice via their GP system supplier. The exact method, form and timing of this invitation will vary by system supplier. However, the invitation will include instructions on how to comply with the DPN, this is a simple and straight forward task. GP system suppliers will commence extractions for individual General Practices who have responded to their system supplier to confirm they are complying with this Notice and provide this data to NHS Digital seven weeks from the date of issue of this Notice, from 1st July 2021. There are two collections within General Practice Data for Planning and Research:

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

1. Patient-level collection – a pseudonymised, full point in time (snapshot) extract of GP medical records containing all content within the scope as set out in Form of the collection above (excluding appointment data).
2. Appointment collection – an incremental extract of appointment data (including appointment management and administration data), with pseudonymised linkage where applicable, containing changed appointment data since the last collection. This will eventually replace the current GPES extraction of appointment data under the General Practice Appointments Data Collection in Support of Winter Pressures 2017 Direction and the GP Appointments Data Collection in Support of Winter Pressures Version 2: Categorisation Data Provision Notice

Scheduling and frequency of patient level and appointment data extraction will be configured as specified to GP system suppliers within the extraction schedule[7]. Patient level extracts and appointment extracts are segregated and flow to NHS Digital's Data Processing Services (DPS) as two distinct flows.

After a GP Practice has confirmed its compliance with the Data Provision Notice to its GP system supplier, all extraction and submission is automated by your GP system supplier, as your processor on your behalf. No further intervention or action is required by General Practices. NHS Digital has worked with your GP system suppliers to ensure this collection does not have any impact on the smooth running of your systems.

Data linkage and pseudonymisation

Data linkage can take place internally within NHS Digital in accordance with and for the purposes set out in the Direction and in relation to dissemination, if approved by DARS as part of a data release.

To link data, the different data assets will be placed in a shared area within NHS Digital's Data Processing Services (DPS). A common pseudonym is created to link the different data assets (usually a pseudonymised NHS Number). This will allow the necessary linkage to take place without the need for directly identifiable patient information to be exposed.

Further detail on pseudonymisation is provided in Appendix A.

Re-identification of the pseudonymised data will only take place where the data is required in directly identifiable form and there is a legal basis to permit this, for example, with express patient consent to share identifiable data with a researcher, or where this is strictly necessary for internal analysis carried out by NHS Digital, which has been subject to independent oversight by IGARD and representatives of the BMA and the RCGP (the Professional Advisory Group). Release of identifiable data will only take place following approval of a specific data access request through the Data Access Request Service subject to independent oversight by IGARD and consultation with the Professional Advisory Group. This will also require compliance with the common law duty of confidentiality, for example where there is approval under section 251 of the National Health Service Act 2006, following support from the Confidentiality Advisory Group, or express patient consent.

[7] This will be made available via https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

More information about the data sets and collections that NHS Digital holds and that may be used for linkage can be found on the NHS Digital Data Collections and Data Sets webpage.

Period of the collection

The data collection is due to commence seven weeks from the date of issue of this Notice on 1st July 2021 from GP Practices who have complied with this Notice.

The Directions will be reviewed on a periodic basis and the data set will be subject to re-examination on a three-yearly basis by the board responsible for the national governance of data sets and data collections (currently the Data Alliance Partnership Board[68] (DAPB)).

For the duration of the collection there will be ongoing review of data requests and use to assure the necessity and proportionality of data collection. See Appendix A for details.

The Collected Data will be kept for eight years from the expiry of the Direction for legal record keeping reasons. This retention period is in line with the NHS Digital Records Management Policy[9] and the Records Management Code of Practice for Health and Social Care 2016.

Burden of the collection

Steps taken by NHS Digital to minimise the burden of collection

NHS Digital has sought to minimise the burden on General Practice by using an 'extract once, use many' approach, where data from a single extraction can be used for multiple approved purposes.

One aim of this data collection is to reduce the burden for General Practice in controlling patient data and maintaining compliance with relevant Data Protection legislation.

In seeking to minimise the burden it imposes on others, in line with sections 253(2)(a) and 265(3) of the Health and Social Care Act 2012, NHS Digital has an assessment process to validate and challenge the level of burden incurred through introducing new information standards, collections and extractions.

This assurance is carried out by the NHS Digital's Data Standards and Assurance Service (DSAS) who assure burden assessment evidence provided as part of the overarching Data Alliance Partnership Board (DAPB) process. The DAPB, acting under authority of the Secretary of State, oversees the development, assurance and acceptance of information standards, data collections and data extractions for the health and social care system in England.

NHS Digital has also sought to mitigate the risk of additional burden to GP Practices in relation to compliance with GPDR, through providing a GP Privacy Notice which GP Practices can link to and providing supporting information and resources to GPs which they can use to support patient awareness raising, including a patient animation. NHS Digital has also provided a Type 1 opt-out form which patients can download from links in the GP Privacy

[8] In November 2020, the Data Coordination Board (DCB) was replaced by a new, cross organisational Data Alliance Partnership. The strategic Board of the Data Alliance Partnership, the DAPB, approves data collections, extractions, and information standards for health and social care in England. Queries about the DAPB should be directed to the DAPB secretariat at dataalliance.partnership@nhsx.nhs.uk

[9] https://digital.nhs.uk/about-nhs-digital/our-work/keeping-patient-data-safe/gdpr/gdpr-register

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Notice and the NHS Digital Transparency Notice to send to their GP Practice easily to exercise a Type 1 opt-out. NHS Digital realises that there may be additional burden on General Practice staff to administer any additional requests to register Type 1 Opt-outs. We have tried to reduce the burden as much as possible by providing patients with a Form they can use which contains all the relevant information and by including the relevant codes GP Practices need to use to register the opt-out on the Form. We cannot however estimate how many Type 1 opt-out requests there may be.

Detailed burden assessment findings

A full burden assessment will be undertaken as part of the full assurance process for the planned information standard that will be undertaken by the Data Alliance Partnership Board, as required before publication of the information standard. NHS Digital will be starting the process to develop the Information Standard shortly and this is planned for publication later in 2021.

Persons consulted

NHS Digital has, as required under section 258 of the 2012 Act, consulted with organisations and stakeholder groups. In particular NHS Digital has engaged in detail with representatives of General (medical) Practitioners (GPs) through the British Medical Association (BMA) and Royal College of General Practitioners (RCGP) for the three years leading up to the issue of the Directions in 2021.

Organisations and stakeholder groups consulted with include:

- The British Medical Association (BMA)
- The Royal College of General Practitioners (RCGP)
- The BMA and RCGP Joint GP IT Committee
- The Department of Health and Social Care, as directing organisation
- NHSx
- NHS England and NHS Improvement
- The National Data Guardian for Health and Social Care
- Independent Group Advising on the Release of Data (IGARD)
- Clinical Commissioning Groups, Commissioning Support Units and Data Services for Commissioners Regional Offices
- Public Health England
- Research bodies including, Health Data Research UK, Healthcare Quality Improvement Partnership, Genomics England, various university institutions
- 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 and NHS Improvement, Public Health England, Care Quality

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Commission, Local Government Association, Health Education England, Health Research Authority, Association of Directors of Adult Social Services and NHS Digital

- 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
- GP system suppliers contracted to supply systems via the GP IT Futures Lot 1 Framework Agreement
- MedConfidential.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Appendix A – Data Minimisation, Additional Protections and Pseudonymisation

Data minimisation and additional protections

Following discussions with GP professional bodies, the British Medical Association (BMA) and the Royal College of General Practitioners (RCGP), NHS Digital has identified data minimisation principles as listed below. This is to ensure that NHS Digital only extracts and holds the data that it needs. The fundamental principles and approaches adopted in the data minimisation are:

1. GP system suppliers will pseudonymise NHS Number, local patient ID, postcode, date of birth and date of death before flowing the data to NHS Digital (see 'Pseudonymisation' below for detail)
2. GP system suppliers will create and append postcode sector, Lower Layer Super Output Area, week/year of death and week/year of birth before flowing the data to NHS Digital. This reduces unnecessary processing by making these separately available for granular analysis where legal and approved
3. Only structured and clinically coded data will be collected (free text, images and documents will not be collected)
4. Legally restricted codes for Gender Recognition and Human Fertilisation and Embryology[10] will not be collected
5. Certain data items in certain categories will be subject to limitations on the age of the data and will not be collected. Such limitations will apply where historic data is not required to meet existing uses due to its age (rolling 10-year limit on all medications, appointments and referrals)
6. Data will be collected for all patients registered with a General Practice including children and Temporary Registrations
7. The Extract ONLY returns codes that have been recorded within the Patient Record. Codes collected in the data collection will be subject to ongoing audit of actual clinical code use in order to compare what data is approved for specific uses against what data NHS Digital is collecting. Where there is clear evidence that clinical codes are being collected which are not required, we will delete this from the codes that are being collected. This audit process will happen on an annual basis starting from 1st July 2022. Full details of this audit process will be determined and published in this calendar year.

[10] GP system suppliers are instructed to remove record entries with codes present in a configurable filter of SNOMED CT Concept Ids. The application of the filter is to provide a capability to remove code concepts or apply data minimisation rules to the extract in response to legal or policy changes.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Data for Direct Care (by exception)

As outlined within the 'Purpose of the Collection' there are exceptional circumstances, where despite being designed for secondary use, the need to use the data for direct care[11] purposes will be necessary and compelling.

Considerations for this exceptional use will take into account patient benefits, burden to stakeholders, expediency and the risks of its use for the relevant direct care purpose. The pandemic has provided the most recent examples of these types of circumstances where data NHS Digital has collected for secondary use purposes has been used to support direct care e.g. to create the Shielded Patient List and to identify additional patients who were considered clinically extremely vulnerable to prioritise them for COVID-19 vaccination.

In such exceptional circumstances, the data collected via this route may be used for direct care purposes, but this will require additional clinical assurance and mitigations to be put in place by NHS Digital to address any issues which result from the data collection not having been designed for this purpose. Use of data for direct care purposes would follow clinical governance and safety according to current standards (Currently DCB 0129/0160). This would be in addition to internal approval processes outlined in this document which will require Executive level clinical sign off including from the NHS Digital Caldicott Guardian. The BMA and RCGP will also be informed of any such use. All transparency of use as outlined in this document will still be adhered to.

Currently no such direct care uses are planned, however, in the context of the COVID-19 pandemic and public health protection it is reasonable to assume that such scenarios will emerge and NHS Digital wishes to be transparent about this.

Pseudonymisation

The data collection consists of personal data (as defined by GDPR) which includes patient-linked General Practice appointment data. The data collection is personal data because NHS Digital will have the potential to re-identify this pseudonymised data in accordance with the appropriate approvals through the pseudonymisation software. NHS Digital will use its national pseudonymisation tool to provide this additional security measure and protection to this General Practice data.

Pseudonymisation at source is an additional protection that we have implemented at the request of the GP Profession. Pseudonymisation of the data occurs within your GP IT system before the data is shared with NHS Digital. The GP system supplier uses the national pseudonymisation tool to request tokenisation rules and encryption keys to de-identify the data at source with a transit token.

The data remains pseudonymised during processing by NHS Digital and during linkage to other data sets by NHS Digital. As a result of the pseudonymisation process, the following

[11] Definitions of direct care from 2nd Caldicott Review. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf "A clinical, social or public health activity concerned with the prevention, investigation and treatment of illness and the alleviation of suffering of an identified individual. It includes supporting individuals' ability to function and improve their participation in life and society. It includes the assurance of safe and high-quality care and treatment through local audit (identified patient safety), the management of untoward or adverse incidents."

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

personal identifiers will be pseudonymised at source and subsequently processed in this protected form:

- NHS Number
- local patient ID
- postcode
- date of birth
- date of death

Each data recipient receives a different set of pseudonyms, and the data access process re-tokenises the NHS Digital pseudonyms to the recipient's pseudonyms. Identifiers that are re-tokenised are:

- NHS Number
- Local Patient Identifier

Date of Birth, Date of Death and Postcode pseudonyms are not disseminated. These can be provided in their minimised form (e.g. postcode as Lower Layer Super Output Area (LSOA), week/year of death and week/year of birth).

NHS Digital protects people's privacy by:

1. Only using secure standard encryption algorithms. Protection comes from the fact that the encryption uses securely stored, secret keys, not from the use of secret algorithms.
2. The use of randomised encryption means that even if source data is breached, the sources cannot be linked without access to the secure private key held by NHS Digital.
3. The use of two stages of encryption means that data can be linked without ever going back to the original identifier.
4. Original identifiers are not passed into or stored in the De-ID Service of NHS Digital; they are encrypted at source by GP system suppliers on behalf of GP Practices prior to submission.
5. Original identifiers are not shared with data recipients. In fact, even encrypted identifiers (Root IDs) are not shared, so there is no risk of them being decrypted. Instead, a random pseudonym is shared.

Getting back to the original identifier would require breaching three systems: gaining access to the shared data; gaining access to the mapping of pseudonyms to encrypted identifiers (which is stored in a secure environment separate from the data); and gaining access to the secure private key, which is held by NHS Digital, used to encrypt the original identifier.

When access to data and/or linkage is required by NHS Digital for purposes set out in the Direction, its Internal Analysis Approval Process will apply to manage access. This process requires the analyst to discuss and agree the work with the Information Asset Owner and then submit details for review by NHS Digital's Information Governance and Legal teams. They review and confirm NHS Digital's legal basis for the work and compliance with GDPR and the common law duty of confidentiality. This is often an iterative process to refine the analysis and data required. Once the IG and Legal teams are satisfied approval is sought from the Executive Director of Privacy, Transparency and Ethics, Executive Director of Data Services and Caldicott Guardian. As described above, if re-identification is required then discussion with Professional Advisory Group (PAG) and a recommendation from IGARD would also be required.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

We expect that all regular and planned internal analytics will take place on pseudonymised data. It is possible that there could be rare circumstances where NHS Digital staff may require GP data to be re-identified. An example would be if a significant data quality issue occurred that could not be resolved using pseudonymised data. In this event, the analytics team will be required to submit this request to PAG/IGARD for consideration ahead of SIRO, Clinical and IAO sign off.

NHS Digital recognises and manages the increased risk of re-identification when combining different data sets, through a combination of technical, security, contractual and organisational controls. Internal organisational controls include the Internal Analysis Approval Process, mandatory staff training, national security vetting where appropriate for identifiable data and other defined processes for internal users. Technical and security controls include encryption of data at rest and in transit and multi-factor authentication and role-based access control within the Data Processing Services and the Data Access Environment. Contractual controls for external users include Data Sharing Agreements, Data Sharing Framework contracts and DAE End User Access Agreements.

When required and if approved, the pseudonymisation tool is able to re-identify the relevant data items. The necessity, purpose and legal basis for identifiable data to be accessed must be set out and justified within a DARS application. The re-identification process would only be performed upon the approval of the application where it is strictly necessary to achieve the purpose. The approval process involves seeking the views of the Professional Advisory Group (comprising of representatives of the BMA and RCGP) and recommendation by IGARD.

Pseudonymisation tool

Pseudonymisation applied by the GP system suppliers on behalf of GP Practices at source and the technology to provide linkable data without re-identification is provided by a third-party pseudonymisation tool. The supplier of this tool does not have access to the data.

The pseudonymisation tool has been assessed by the NCC Group, who are accredited by the Government Procurement Service to provide products and services to public sector organisations and provide the National Cyber Security Centre's (NCSC - formerly CESG) Tailored Assurance Service (CTAS). NCC Group has performed code review and security audits for the tool and has assessed that the product uses proven cryptographic implementations. Its underlying technology uses Homomorphic encryption which is a form of de-identification and encryption which allows the generation of pseudonymised ID for data linkage without the need for re- identification.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Appendix B – Data Model

The data from GP medical records that will be included in the General Practice Data for Planning and Research data collection can be viewed in the Data Model which is a high-level diagram showing the data items grouped in a relational table format.

The Data Model can be viewed on NHS Digital's website:

https://nhs-prod.global.ssl.fastly.net/binaries/content/assets/website-assets/corporate-information/directions-and-data-provision-notices/data-provision-notices/general-practice-data-for-planning-and-research-dpn-appendix-b-data-model.pdf

Additionally the full data set is specified in Annex B of the General Practice Data for Planning and Research Directions 2021

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

Appendix C – NHS Digital dissemination of General Practice Data

Under section 261 of the 2012 Act, information that NHS Digital obtains by complying with a Direction under section 254 may be disseminated or shared (by way of secure access) as appropriate. Data applicants will need to demonstrate through the NHS Digital Data Access Request Services (DARS) assessment process that they have a lawful basis to access and process the data. Where identifiable patient level data is requested, the data applicant will need a legal basis to process this data without breaching the common law duty of confidentiality. This may include express patient consent or an approval under section 251 of the National Health Service Act 2006 and Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002, following support from the Confidentiality Advisory Group, for example in the case of certain research and clinical trials. Use of data for research purpose will also require a Research Ethics Committee approval.

Requests will be assessed by DARS against specific criteria underpinned by information governance assessment standards and will be subject to oversight by the Independent Group Advising on the Release of Data (IGARD). The DARS process is robust and well-established, and consists of enquiry, triage, review, independent oversight through IGARD, approval, access, audit and destruction phases. The DARS process will apply appropriate additional scrutiny to any data release requiring re-identified data, which will also involve oversight by IGARD and the Professional Advisory Group (representatives of the BMA and RCGP).

Data released through DARS will where possible be provided to data recipients through a secure data access environment within NHS Digital infrastructure, or where the approved requirements of the recipient cannot be met this way, and where the recipient has met relevant DARS standards for dissemination, including but not limited to meeting the DARS security standards, through a direct dissemination of a copy of the relevant minimised data by means of a data extract.

All data approved for access or release through DARS is subject to data sharing agreements between NHS Digital and the Controller requesting the data. More detail on the DARS process, standards and the data sharing agreements used are available here. Details on the IGARD review and oversight process is here. For more information including about IGARD's constitution and its assurance and oversight role, please see its Terms of Reference.

NHS Digital's responsibility for the data does not stop following access being granted or an extract being disseminated. Audits are carried out and sanctions are imposed for any organisation deemed to have breached the Data Sharing Agreement (DSA). These include:

- terminating the DSA and revoking access to the data
- requiring the data to be securely destroyed and confirmation provided in the form of a data destruction notice signed by the Data Protection Officer or other suitable individual
- reporting any potential personal data breach (as defined in GDPR) to the organisation's Data Protection Officer and in certain cases to the Information Commissioner's Office.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.

In the event that IGARD does not recommend approval of a request for access to this data and NHS Digital disagreed with that recommendation, NHS Digital would generally seek guidance from the National Data Guardian and/or the Confidentiality Advisory Group before disseminating any data.

The National Data Opt-Out will be applied by NHS Digital in line with National Data Opt-out Operational Policy Guidance. For more information on the application of the National Data Opt-out and exemptions to this, see Chapters 5 and 6 of this Guidance.

For further information

www.digital.nhs.uk

0300 303 5678

enquiries@nhsdigital.nhs.uk

Under the Open Government Licence you are encouraged to use and re-use the publicly accessible information in this notice free of charge. Re-use includes copying, issuing copies to the public, publishing, broadcasting and translating into other languages and its subsequent use in commercial or non-commercial enterprise.

Copyright © 2021 Health and Social Care Information Centre.
The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital.