Document filename:	**Data Protection Impact Assessment**		
Directorate / Programme	**GPES Data for Pandemic Planning and Research (COVID-19)**		
Document Reference	IAR0000886		
Information Asset Owner	*Dave Roberts*	Version	**0.9**
Author	Tess Morley, Richard Birmingham and Claire Kwon	Version issue date	**10/112020**

DATA PROTECTION IMPACT ASSESSMENT: GPES Data for Pandemic Planning and Research (COVID-19)

Document Management

Revision History

Version	Date	Summary of Changes
0.1	14/05/2020	Initial Draft
0.2	15/05/2020	Overall revision based on the updated DPN and feedback from Peter Short, Jackie Gray, Andrew Thorne-Marsh and Susannah Strong; Addition of risks and mitigations in Part C and Part E
0.3	20/05/2020	Alignment with published DPN; Overall revisions in line with feedback from reviewers
0.4	05/06/2020	Overall revisions in line with feedback from reviewers
0.5	06/06/2020	Overall revisions in line with feedback from reviewers
0.6	09/06/2020	Update following review by Arjun Dhillon, Dave Roberts, Garry Coleman
0.6a	07/07/2020	Further updates following reviews by Dave Roberts, Garry Coleman, Richard Birmingham and others
0.7	11/08/2020	Further update with feedback following review by Jackie Gray
0.8	15/09/2020	Overall revisions in line with feedback from reviewers
0,8	10/11/2020	Table of applicable rights added and legal review

Reviewers

This document must be reviewed by the following people:

Reviewer name	Title / Responsibility	Date	Version
Susannah Strong	Senior Communications Manager	07/05/2020	0.1
Andrew Thorne-Marsh	Programme Manager	11/05/2020	0.1
Jackie Gray	Executive Director of Information Governance	11/05/2020	0.1
Peter Short	Clinical Lead	13/05/2020	0.1
Carol Lodge	Business and Operational Delivery Manager	19/05/2020	0.2
Eva Simmonds	Primary Care Technology Programme Head	19/05/2020	0.2
Kathryn Salt	Information Analysis Lead Manager	19/05/2020	0.2

Kevin Willis	Data Protection Officer NHS Digital	20/05/2020	0.2
Jackie Gray	Executive Director of Information Governance	21/05/2020	0.2
Garry Coleman	Associate Director, Data Access	02/06/2020	0.3
Deborah Raven	Knowledge Manager, Data Standards Assurance	04/06/2020	0.3
Jackie Gray	Executive Director of Information Governance	27/05/2020	0.3
Roy Taylor	Solutions Architect	02/06/2020	0.3
Andrew Thorne-Marsh	Programme Manager	02/06/2020	0.3
Susannah Strong	Senior Communications Manager	08/06/2020	0.5
Arjun Dhillon	Caldicott Guardian	08/06/02020	0.6
Dave Roberts	Information Asset Owner	08/06/2020	0.6
Garry Coleman	Associate Director, Data Access	08/06/2020	0.6
MedConfidential		11/06/2020	0.6
Andrew Thorne-Marsh	Programme Manager		0.6
Deborah Raven	Knowledge Manager, Data Standards Assurance		0.6
Jackie Gray	Executive Director of Information Governance		0.6
Tim Gentry	Chief Technology Programme Director	23/06/2020	0.6
Roy Taylor	Solutions Architect	23/06/2020	0.6
Jackie Gray	Executive Director of Information Governance	16/08/2020	0.7
Garry Coleman	Associate Director, Data Access	20/09/2020	0.8
Dave Roberts	Information Asset Owner	20/09/2020	0.8
Stuart Crook	NHS Digital Lawyer	01/11/2020	0.8
Richard Birmingham	NHS Digital, Privacy, Transparency and Ethics	10/11/2020	0.9

Approved by

This document must be approved by the following people:

Name	Title / Responsibility	Date	Version
Arjun Dhillon	Caldicott Guardian		
Jackie Gray	Executive Director of Information Governance		
Kevin Willis	Data Protection Officer		
Dave Roberts	Information Asset Owner		

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 (e.g. paper, email attachment), are considered to have passed out of control and should be checked for currency and validity

Contents

BACKGROUND AND CONTEXT

Background and Context

NHS Digital was established under the Health and Social Care Act 2012 and is the national statutory safe haven of patient data in England with statutory powers to collect and analyse information, including confidential patient information, publish anonymous information and disseminate information, including confidential patient information, to those organisations with a lawful basis to process it.

NHS Digital has been requested by representatives of the GP Profession to collect data from general practices in England for COVID-19 pandemic planning and research purposes: GPES Data for Pandemic Planning and Research (COVID-19) – (**GDPPR**). This is needed to respond to the intense demand for General Practice data to be shared in support of vital planning and research for COVID-19 purposes, and to relieve the growing burden and responsibility on general practices.

The Secretary of State for Health and Social Care has directed NHS Digital under section 254 of the Health and Social Care Act 2012 (the **2012 Act**) by the COVID-19 Public Health Directions 2020 17 March 2020 (as amended) (the **COVID-19 Direction**) to collect, process and analyse patient data (the **Collected Data**) extracted from patient health records of general practices in England. This direction is referred to in this document as the COVID-19 Direction.

NHS Digital issued a Data Provision Notice (**DPN**) to general practices in England on 14 May 2020 pursuant to Section 259(1)(a) and 259(5) of the 2012 Act, requiring them to provide the data in the form and manner specified in the DPN.

The England General Practitioners Committee of the British Medical Association (**BMA**) and the Royal College of General Practitioners (**RCGP**) have also requested that NHS Digital collect clinically relevant data centrally for the purposes outlined in the DPN and they have requested support from General Practitioners for these data collections.

It is a requirement of the Joint GP IT Committee (**JGPITC**) (the BMA and RCGP) that all requests by organisations to access and use this data to support the COVID-19 response will need to be made via the NHSX Single Triage Service (also known as the Single Point of Contact or SPOC). The SPOC will triage and prioritise these requests and refer appropriate requests on to the NHS Digital Data Access Request Service (**DARS**).

NHS Digital will involve representatives of the BMA and the RCGP in reviewing requests for access to the data. An outline of the process for this agreed with the BMA and the RCGP is published here. Data applicants will need to demonstrate they have a lawful basis to access the data for COVID-19 purposes. Requests by organisations to access record level data from this collection will also be subject to independent scrutiny by the Independent Group Advising on the Release of Data (**IGARD**).

The initial GDPPR extract will consist of patient demographic information and coded medical information (as per the business rules https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof#other-extracts) as a snapshot in

time when the first extract is undertaken. A snapshot in this context means data recorded up to the date the extract is taken, looking back through the full history of the relevant parts of the patient record stored within their GP system. Thereafter subsequent fortnightly extracts will then be taken. The fortnightly extracts will ask for the same data items (patient demographics and coded medical information) and snapshot as defined in the initial extract but from a more specific group of patients, namely any who meet at least one of the criteria below. This group of patients are described as below;

1. patients who have recently registered at a GP practice in the two weeks up to and including the reporting period end date.

2. patients who have any codes relevant to pandemic planning and research recorded in the month up to and including the reporting period end date

3. patients who have any codes relevant to pandemic planning and research and whose date of death is in the month up to and including the reporting period end date

In summary:

We will collect updates where:

- patients register at a new practice
- journals are added
    - journals are added and removed in between reporting periods
    - patients have died

We will not collect updates where:

- only changes made in the patient section of the record
- only journals are removed
- only contents of journals are changed
- patients are deleted from practice registers

The subsequent fortnightly extraction will then continue until the expiry of the COVID-19 Direction. This is currently 31 March 2022 but will be reviewed in September 2020 and every six months thereafter. The frequency of the data collection may change in response to demand.

Upon the expiry of the COVID-19 Direction the Collected Data will no longer be analysed by NHS Digital, unless there are replacement Directions which would permit this. It may continue to be shared with and used for COVID-19 purposes by those who have a legal basis to process it for COVID-19 purposes.

A wide range of organisations are expected to access the Collected Data, including CCGs for COVID-19 planning and commissioning purposes and researchers for a range of COVID-19 research purposes. NHS Digital will use its powers under the COVID-19 Directions and section 261 of the Health and Social Care Act 2012 to share information and under the notice issued to it under Regulation 3(4) of the Health Services (Control of Patient Information) Regulations 2002 (NHSD COPI Notice). Other organisations can also access and process data comprised in the

Collected Data for COVID-19 purposes under notices issues to them under Regulation 3(4) of the Regulations.

The Office for National Statistics may also apply to access the data using their statutory powers under Part 5 of the Digital Economy Act 2017. All applications will be assessed by DARS in line with the process described above.

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.

The GP Practices are the controllers of the Collected Data before it is extracted and shared with NHS Digital. When it has been collected by NHS Digital, NHS Digital becomes the controller of the Collected Data. The collection by NHS Digital of this Collected Data is considered to require a DPIA to be carried out by NHS Digital. NHS Digital has therefore prepared this document as its DPIA to satisfy its own compliance requirements as a controller of the Collected Data under the COVID-19 Direction.

SECTION 1 DPIA SCREENING PROCESS

1. Initial DPIA screening questions for GPES Data for Pandemic Planning and Research (COVID-19)

The following questions were used to determine whether GDPPR requires a DPIA. If the answer to any of the following questions is yes, a full DPIA is required.

No.	Question	Yes/ No	Explanation
1.	Does the proposal involve any evaluation or scoring including profiling & predicting using information about a p4erson?	No	The GDPPR extract does not involve any evaluation or scoring including profiling & predicting using information about a person. When considering dissemination of the data, currently there are no use cases for this, but such applications may be considered if the purpose for any profiling using the data is for a COVID-19 purpose.
2.	Does the proposal involve any automated decision making which has a legal or similar legal effect e.g. whether to employ an individual, grant them a loan or offer medical insurance?	No	The GDPPR extract does not involve any automated decision making which has a legal or similar legal effect. When considering dissemination of the data, currently there are no use cases for this, but such applications may be considered if the purpose for any automated decision making using the data is for a COVID-19 purpose.

No.	Question	Yes/ No	Explanation
3.	Does the proposal involve any systematic monitoring: processing used to observe, monitor or control individuals, including data collected through networks e.g. employees' activities, including the monitoring of the employees' work station, internet activity; monitoring of wellness, fitness and health data via wearable devices; closed circuit television; connected devices e.g. smart meters, smart cars, home automation; includes internet tracking and profiling for behavioural advertisement?	No	The only data collected is that which has been entered onto the GP patient record. Currently there are no use cases for this, but such applications may be considered if the purpose of monitoring is for a COVID-19 purpose.
4.	Does the proposal involve any sensitive information or information of a highly personal nature e.g. health?	Yes	Data shared by a GP Practice will include data about all relevant patients registered with a Practice as set out in the DPN, including data held in their GP patient records about their health and medical diagnosis. It will also include the ethnicity special category of data. For more information on the data to be collected see the Data Provision Notice.
5.	Does the proposal involve data processed on a large scale? Large scale is not defined but should consider: A) The number of data subjects, either as a specific number or as a proportion of the relevant population. B) The volume of data and/or the range of different data items processed.	Yes	The data collection is expected to cover more than 50% of patients in the participating GP Practices in England. It is expected there will be many requests for access to and dissemination of relevant data from this collection by third party organisations for COVID-19 planning and research purposes, which may include linkage of the data with other data sets.

No.	Question	Yes/ No	Explanation
	C) The duration, or performance of the data processing activity. D) The geographical extent of the processing activity. Processing of patient data in the regular course of business by a hospital would be classed as “large scale” while processing of patient data by an individual physician would not.		However, where such dissemination occurs, the data being disseminated must be minimised such that it is restricted to that which may be reasonably required in order to meet the defined purpose. Such minimisation is documented within the application that is considered through the DARS process and IGARD.
6.	Does the proposal involve any matching or combining of data sets? i.e. matching two or more data processing operations performed for different purposes in a way that would exceed the reasonable expectations of an individual.	Yes	NHS Digital will perform analysis and linkage of the collected data to other data sets that NHS Digital holds where required for COVID-19 purposes under the COVID-19 Direction. This might for example include data on hospital attendance, demographic data, or other COVID-19 related datasets such as Testing data. Applicants seeking access to the data are also likely to ask NHS Digital to link the data to other data we hold and may want to link the data to other data they hold for COVID-19 purposes. NHS Digital has published a GPES Transparency Notice to describe how the data will be used and will publish details of disseminations in its Data Release Register
7.	Does the proposal involve any data concerning vulnerable individuals who may be unable to easily consent or oppose the processing, or exercise their rights?	Yes	The data that is collected will include data about patients from patient records including records from children, people with mental health problems and

No.	Question	Yes/ No	Explanation
	This group may include children, employees, mentally ill persons, asylum seekers, or the elderly, patients and cases where an imbalance in the relationship between the position of the individual and the controller (NHS Digital) can be identified.		other vulnerable individuals.
8.	Does the proposal involve any innovative use or applying new technological or organisational solutions e.g. combining use of fingerprint and face recognition for improved physical access control?	No	Currently there are no use cases for this, but such applications may be considered if it was for a COVID-19 Purpose.
9.	Does the proposal involve any processing which in itself 'prevents data subjects from exercising a right or using a service or contract' e.g. determining eligibility based on an individual's circumstances?	No	GDPPR will recognise Type 1 Objections when collecting the data and apply National Opt-Out where appropriate on dissemination.
10	Will tracking on individual's geo-location or behaviour including but not limited to the online environment occur?	No	Not applicable – individual tracking is not relevant in relation to the data collection.
11	Will Invisible Processing occur?	Yes	NHS Digital may carry out analysis and linkage of the data where permitted under the COVID-19 Direction, although we have published a COVID-19 Transparency Notice explaining how we process data for COVID-19 Purposes, and a specific GPES Transparency Notice for this collection.

No.	Question	Yes/ No	Explanation
12	Are any of the following risks envisaged as a result of this project? For example: • risk of identity theft; • risk of financial loss; • risk of mental distress; • risk of reputational damage • risk of physical harm to individuals.	Yes	There are risks of mental distress and perceived reputational damage as a result of the data being collected, processed and shared. There is a risk that there is a breach of security in the transmission of the data to NHS Digital which results in the unauthorised disclosure of personal data. For information on how these risks will be mitigated, see Section 2 Part E below.

2. Conclusion

High risk processing activity has been identified through carrying out the Screening Questionnaire above in relation to GDPPR. This means that a DPIA needs to be carried out under Article 35 of GDPR. Mitigations for the risks identified above are included in Part E of Section 2 of this document.

SECTION 2 – FULL DPIA PROCESS

This DPIA should be kept up to date as this data collection is designed, developed and implemented. It should be reviewed whenever there is a significant change to processing activity.

PART A: PURPOSE OF PROCESSING

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

Data may be analysed and linked by NHS Digital to other data held by NHS Digital for COVID-19 purposes under the COVID-19 Direction. NHS Digital will only make the data available for COVID-19 purposes.

Other organisations can also make requests to access and use this data via the NHSX Single Triage Service (also known as the Single Point of Contact or SPOC) to request access to health and care data in order to support the COVID-19 response. NHSX will triage and prioritise these requests and refer appropriate requests on to the NHS Digital DARS service for assessment and fulfilment where approved.

Requests by organisations to access record level data from this collection will also be subject to Independent Group Advising on the Release of Data (**IGARD**) consideration. Data applicants will need to demonstrate they have a lawful basis to access and process the data for COVID-19 purposes.

Data may be linked to other data held by NHS Digital at the request of applicants, or once disseminated linked to other data held by those applicants for COVID-19 purposes. Applications will be assessed through the process described above.

Data to be shared by NHS Digital will also be subject to data minimisation rules, restricting the data shared to that which may be reasonably expected to be required for the approved purpose. For example, data shared with a clinical trial would be restricted to including those people who are participating within the clinical trial. Additionally, not all detail on an individual need be shared. For example, if a research study is considering the impact of COVID-19 on the particular disease, then only data relating directly to that study would be shared, and only for people who might quality to be included within that study.

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 participating GP Practices in England every fortnight. GP IT System Suppliers will extract data already held in GP Practice patient record systems and transfer this data to NHS Digital using the established General Practice Extraction Service (GPES) tool. The data provided will therefore be data which 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.

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 and planning purposes. A central collection, with NHS Digital managing all dissemination of the data 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. It will also reduce compliance burden and risk for General Practice associated with sharing data and complying with the terms of the general legal notice issued under the Health Service (Control of Patient Information Regulations) 2002 (COPI), which applies to General Practices.

Consequently, data required by the Government, scientists, researchers and those who plan and run health and care and other public services 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.

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

The Government, scientists, the health and social care organisations and researchers need access to this vital data for a range of COVID-19 purposes, to help plan, monitor and manage the national response to the COVID-19 pandemic, 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, information about incidents of patient exposure to COVID-19 and the management of patients with or at risk of COVID-19 including: locating, contacting, screening, flagging and monitoring such patients and collecting information about and providing services in relation to testing, diagnosis, self-isolation, fitness to work, treatment, medical and social interventions and recovery from COVID-19
- understanding information about patient access to health services and adult social care services as a direct or indirect result of COVID-19, and the availability and capacity of those services
- monitoring and managing the response to COVID-19 by health and social care bodies and the Government including providing information to the public about COVID-19 and its effectiveness and information about capacity, medicines, equipment, supplies, services and the workforce within the health services and adult social care services

- delivering services to patients, clinicians, health and adult social care services workforce, and the public about and in connection with COVID-19, including the provision of information, fit notes and the provision of health care and adult social care services; and
- research and planning in relation to COVID-19, including COVID-19 related clinical trials.

This will be a significant reduction of 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. It will also reduce compliance burden and risk for General Practice associated with sharing data and complying with the terms of the general legal notice issued under COPI, which applies to General Practices.

Further information about this data collection can be found at the following locations on NHS Digital's website:

https://digital.nhs.uk/coronavirus/gpes-data-for-pandemic-planning-and-research

PART B: CONSULTATION

1. Outline any work that you have undertaken to date, to engage with GPs patients, the public and other stakeholder groups in relation to GPES Data for Pandemic Planning and Research (COVID-19)?

NHS Digital, in collaboration with the Department of Health and Social Care (on behalf of the Secretary of State) aims to build informed and clear public support for the use of patient data for pandemic planning and research in relation to COVID-19.

Stakeholder feedback and advice also informed the development of other critical work areas for GDPPR, including:

- Information governance – Ensuring data is protected and processed in line with national policy and legislation.
- Business processes – Developed and implemented in line with the needs of our stakeholders.
- Technical architecture – Ensuring the collection and processing of data is aligned to information governance, business processes and the Data Processing Services (DPS) requirements.
- We have consulted with a number of stakeholders including those who represent patient and public interest in data – for example the Office of the National Data Guardian and Healthwatch England (see table below).

Additional feedback on our transparency for the extract and specifically on the GDPPR GP template transparency notice has been sought from patient participation groups at HDRUK and Genomics England. Looking ahead we will continue to seek views of and engage with these and other patient and public interest groups in the interests of full transparency and awareness.

Key Stakeholder Groups and Consultation Methods Used:

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
Representatives of GPs in the BMA and RCGP Note: chairs' of JGPITC approached NHS Digital to ask for support in managing increasing number of data requests during COVID-19 so as to reduce burden on	The RCGP is the professional body for general (medical) practitioners in the UK. The RCGP represents and supports GPs on key issues including licensing,	Engagement and involvement has been via virtual meetings, calls, letters and email correspondence with the	Requested a tactical solution from NHS Digital to meet the demand for GP data in support of urgent care planning and research directly related to COVID-19 and to relieve the	NHS Digital responded to the request with a proposal to use the General Practice Extraction Service (GPES) to deliver a data collection from General Practice at scale and pace on a 'tactical' basis to support the COVID-19 response. The

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
GPs so they can better focus on patient care.	education, training, research and clinical standards. The BMA represents the interests of GPs and patients. across a range of activities and special interests. Both have an interest in changes to working practices that impact the GP community, and ensuring that the new processes are not deemed excessive and/or impracticable to implement/support. Both have an interest in reducing burden on the GP community.	joint chairs of the JGPITC[1] and others NHS Digital Executive Director of Information Governance, Caldicott Guardian, GP clinical lead and Information Asset Owner along with NHS Digital communications team have engaged with these stakeholders to agree end to end governance steps and to ascertain best ways and means of communicating with general practices.	burden/responsibility on General Practice. There must be due diligence regarding any applications for the data that is being collected. Applications must pass the appropriate ethical, legal and Information Governance tests, and involve a review by representatives of RCGP and BMA ahead of data being disseminated. Want to be party to any direct communications approach to GPs.	BMA/RCGP gave its support via JGPITC to this proposal which they have endorsed within the DPN itself. Additional steps in the NHS Digital data access request service process have been agreed to involve RCGP and BMA representatives in the review of all applications for data. Communications plan takes account of role of BMA/RCGP in GP communications.
National Data Guardian	The Office of the National Data Guardian (NDG) is a public body that advises and challenges the health and care system to ensure that citizens' confidential information is safeguarded securely and used properly.	Virtual meetings, calls and email correspondence have been undertaken due to social distancing restrictions. Requested NDG panel review of the proposed tactical solution and	Data must be extracted and disseminated in line with NDG recommendations for the protection of data subject rights and freedoms including data minimisation. Transparency and awareness of this collection is paramount to	NHS Digital accepted all recommendations. Plans for transparency for patients include activity to raise awareness of the use of patient data to support the COVID-19 response. NHS Digital has responded to NDG's questions regarding

[1] The Joint GP IT Committee is a contractually-mandated committee representing the views of GPs from all 4 UK nations and users of all GP systems in discussions relating to the use and management of these systems and GP data created therein.

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
	The NDG represents the voice of the patient and wants to ensure that patients are informed, and that any data extracted and shared is being done so in a secure and compliant manner that protects the rights and interests of individuals.	received formal panel response. Direct discussions with the NHS Digital Executive Director of Information Governance and NDG to agree additional measures as part of the end to end assurance process. Presentation to the NDG Panel by the GDPPR Programme Head.	build/maintain public trust. Concerned over cyber-security risks, noting the increase during the pandemic. The Office of the NDG confirmed written support to the approach with certain criteria to be met.	governance and security to their satisfaction. The NDG panel will continue to be briefed and comment upon communications activities during the collection period.
Information Commissioner's Office (ICO)	The ICO is the UK supervisory authority under the General Data Protection Regulation. Where a controller under GDPR has not been able to mitigate high risks to the rights and freedoms of individuals through the mitigations it has put in place for a project, the controller is required to consult the ICO under Article 36. NHS Digital does not consider it is required to consult the ICO under Article 36 but is consulting with the	Virtual meetings, calls and email correspondence. NHS Digital's Executive Director of Information Governance has been in contact with the appropriate health sector ICO representative to explain the project. A copy of the transparency material was shared, and comments received.	The ICO does not need to be formally consulted in relation to this DPIA as this data collection and dissemination does not reach the threshold where there are high risks to individuals that cannot be mitigated. NHS Digital will however ask the ICO to review and comment on the final version of the DPIA. NHS Digital has asked the ICO to review informally and comment on both the NHSD Transparency Notice and	The ICO has reviewed the transparency notices and suggested improvements which have been actioned. The final draft version of this DPIA will be sent to the ICO for their informal review and feedback

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
	ICO on a voluntary basis for comment and feedback on Transparency Notices and this DPIA.		the template Transparency Notice for GPs.	
NHSX	NHSX brings together teams from DHSC, NHS England and NHS Improvement to improve health and social care through technology. NHSX supports the data collection to provide data for research and planning in relation to COVID-19 which can be shared, including where required under COPI Notices.	Virtual meetings, calls and email correspondence between various representatives. This includes discussions between NHS Digital's Executive Director of Information Governance and the NHSX's Deputy Director, Data and Information Governance Policy and the Chief Executives of NHS Digital and NHSX, and consultation with the NHSX Director for Digital Primary Care. As part of the NHSX COVID-19 Front Door, an application form has been implemented to support requests for GDPPR data from NHSD.	GDPPR needs to align with NHSX approach to data sharing in support of COVID-19. All requests for GDPPR data must be directed to NHSX Single Triage Service for triage before appropriate requests are redirected to DARS. NHSX Director for Digital Primary Care has supported the discussions with the BMA and RCGP.	NHS Digital has worked with NHSX to implement the process for triage and handover of data requests and ensuring suitable communications are in place to direct initial requests to NHSX.
Public Health England (PHE)	PHE is a public body which promotes the health and wellbeing of England.	Email Correspondence with the Consultant Epidemiologist (Real-Time	Has established data flows in place to meet surveillance need but may request access to the GPES extract	PHE is aware they can make an application for this data at any point.

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
	Potential significant user of the data.	Syndromic Surveillance) for National Infection Service and Chief Clinical Information Officer.	relating to COVID-19 to look at wider Public Health implications.	
Clinical Commissioning Groups (CCGs)	CCGs are public bodies responsible for commissioning health services for their local community. Need assurance that any dissemination of data by NHS Digital is scrutinised and upholds patients' rights. Potential significant users of the data.	Virtual meetings, calls and email correspondence with CCG Leads via Data Services for Commissioners CCG Steering Group to discuss the breadth of collection and appropriateness of use in local planning of health services during COVID-19. At a steering group meeting, CCGs expressed an interest in using GP data in relation to COVID-19; 26 CCGs submitted statements setting out a the need for CCGs to have GP data for planning purposes in relation to COVID-19.	CCGs require timely access to the data required to meet local planning of health services and monitoring requirements, including data from General Practice. Concerned as to potential delays and burden introduced by NHSX Single Triage Service.	NHS Digital is working directly with CCGs to agree a set of data required to meet local planning requirements in support of the COVID-19 response. NHS Digital team is a member of the NHSX Single Triage Service review panel and has helped to design a streamlined process for applications to data held by NHS Digital. This will be reviewed frequently.
Health and Social Care Research Bodies	The research bodies represent the interests of researchers who may be commissioned to conduct research in relation to COVID-19.	Virtual meetings, calls and email correspondence with (e.g.) • Professor of Neurology and Clinical Epidemiology, Centre for Medical Informatics,	Whether the collected data will have the right data items and coverage of patients to deliver the required utility in support of research related to COVID-19.	Direct engagement with research body representatives to reach agreed position on the data items and inclusion criteria, balanced against need for delivery at pace and

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
	Several research bodies have approached NHS Digital to discuss their needs for GP data for pandemic related research. Want the data collection to provide data for research and planning in relation to COVID-19.	Usher Institute, University of Edinburgh, • Director, BHF Data Science Centre • Director, HDR UK Scotland • Partnership Director, Health Data Research UK,.	Need for delivery at pace to support urgent research into COVID-19. Whether the collected data will be accessible to researchers. Data needs to be refreshed frequently so that it is up to date as possible (data freshness). Want to be able to link to other data sets where there is a legal and ethical basis, including where NHS numbers (the main method of linkage).	capacity of GPES and system suppliers.
Understanding Patient Data (UPD)	Understanding Patient Data aims to make uses of patient data more visible, understandable and trustworthy, for patients, the public and health professionals. https://understandingpatientdata.org.uk/about-us	Virtual meetings, calls and email correspondence with the Lead of UPD.	Raised concerns of potential confusion for patients due to conflation with Contact Tracing App and NHSE/I COVID-19 data store. Communications approach and materials shared with UPD representatives.	Content and timing of patient and GP facing communications informed by UPD observations. Ongoing relationship to ensure the language and approach to awareness raising is beneficial and is aligned with learning from other areas and studies.

Stakeholder Consulted	Representative Because…	Consultation Method	Views Raised	How NHS Digital is taking views into account
Healthwatch England	Patient representative group Healthwatch England was established as an effective, independent consumer champion for health and social care. It also provides a leadership and support role for the local Healthwatch network.	Programme head met virtually the policy lead of Healthwatch England.	Raised concerns of potential confusion for patients due to conflation with Contact Tracing App and NHSE/I COVID-19 data store. Provided view on the key areas our comms and transparency materials need cover.	NHS Digital reviewed all patient facing materials to ensure transparency, including a clear explanation regarding patients right to opt out.

PART C: THE NATURE OF THE PERSONAL DATA USED

2. Personal data processed

Complete the table below setting out what data will be processed in submitting the Collected Data to NHS Digital for GDPPR:

Type of data	Do you process these data? (yes or no)	If yes, please list the types of data you process
Basic identification personal data (e.g. name, address, email, telephone, etc.)	Yes	• Patient Demographic Data, including NHS Number, Date of Birth, Date of Death, Surname, Forename, Address, Postcode, and Sex. See the Data Provision Notice for more information.
Special category data (e.g. health data, race or ethnic origin, biometric data, , sex life/sexual orientation)	Yes	• Physical / Mental Health or Condition • Ethnic Origin The Project will include sharing personal data that is coded and structured in the following categories from the GP clinical record system: • diagnoses and findings • medications and other prescribed items • investigations, tests and results • treatments and outcomes • vaccinations and immunisations. See the Data Provision Notice for more information.

3. Description of the processing to be carried out

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
1	Nature of Processing Compliance with Lawfulness, Fairness and Transparency Principle	Explain how personal data will be collected (i.e. • Will you be collecting it directly from individuals? • Do you already have the personal data, and if so, where did you get it from? • Did you obtain the personal data from another organisation?)	The Practice collects data from patients via registration forms and through appointments and treating the patient. Reports and test results are received from other medical providers and health care professionals involved in the patient's care which are also included in the patient health record. The data to be shared with NHS Digital is a subset of data which the GP Practice already holds in the patient records systems. The required data will be collected from General Practices' clinical IT systems via the General Practice Extraction Service (GPES). The NHS Digital GP Collections webpage[2] provides further information on this service. The extract will be transferred to NHS Digital Data Processing Services (DPS) platform via Message Exchange for Social Care and Health (MESH).	Risk that more data than is necessary for NHS Digital's purposes will be shared by GP Practice. (Risk 3) Risk that there is insufficient transparency to data subjects of the processing of their personal data for the purposes of the Project. (Risk 2) Risk that National Data Opt-Outs will not be respected. (Risk 4) Risk that Type 1 Opt-Outs will not be respected. (Risk 4)
2	Nature of Processing Compliance with Purpose	How will personal data be used?	Under the COVID-19 Direction, the Collected Data may only be used by NHS Digital and by other organisations for the purposes set out in the COVID-19 Direction[3] and as described in the Background and Context section of this DPIA.	Risk that Collected Data is used for purposes other than as specified in the COVID-19 Direction. (Risk 6,8,14) Risk that more data than is

[2] https://digital.nhs.uk/services/general-practice-gp-collections

[3] https://digital.nhs.uk/binaries/content/assets/website-assets/corporate-information/directions-and-data-provision-notices/secretary-of-state-directions/nhsd190606a-iii-p2_final-sofs-covid-19-directions-nhsd.pdf

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
	Limitation Principle		See also below under 4. Nature of Processing - Compliance with Integrity and Confidentiality Principle for how access is granted to data held. Virtual result sets (or data views) of the data will be developed and provided as security-controlled data assets through NHS Digital's Data Access Environment (DAE), the default method of dissemination for GDPPR data. Data views are pre-existing queries, which provide a user with access to a constrained version of the underlying data. Each data view is a restricted window on the underlying data, not a copy of it. NHS Digital creates specific views of data within the DAE for customers, which reflect what the customer is approved to access in their Data Sharing Agreement, depending on the DARS-authorised access requirements for a particular applicant. Our approach to the creation of Views within DAE may change over time in response to the nature and volume of DARS requests we receive. Any changes will be documented in future versions of this DPIA. Where DARS-authorised requests permit a data file for dissemination, NHS Digital will use the customer-specific view to produce the data file for release via the normal SEFT and MESH processes. Where users access the data within the DAE and where a user is not entitled to see Personal Identifiable Data, the linkage will be done by joining datasets with the appropriate domain pseudonym. Domain pseudonyms will only be generated for the datasets	necessary for NHS Digital COVID-19 purposes will be analysed and linked by NHS Digital. (Risk 6) Risk that more data than is necessary for third-party organisations COVID-19 purposes will be analysed and linked by that organisation (Risk 8)

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			covered by their associated Data Sharing Agreement. Where identifiable GDPPR data is sent to data requesters and can be linked to other datasets that the requester holds, under the permissions of the Data Sharing Agreement, the requester would generally link datasets using NHS Number and Date of Birth. When deidentified, data requesters would use the domain pseudonym to link datasets.	
3	Nature of Processing Compliance with Storage Limitation Principle	Explain how personal data will be stored (e.g. • on which IT platform? • In hard copy? • Third party cloud storage? • and for how long)	The DPN and the GPES business rules provided to suppliers set out the scope of the collection. The GP IT System Suppliers develop the extract in accordance with the Business Rules. Once the extract is developed, GPES will be used to schedule and manage the collection and onward processing of the data into DPS. GPES (General Practice Extraction Service) is an established mechanism to schedule, extract and deliver GP Practice data from GP system supplier clinical systems. For the purposes of the GDPPR collection, it is made up of three key components: • GPDC: The **GP Data Collector** is the solution operated by the NHS Digital's Data Services Alliance team. It will send requests for data to the GP System Supplier solutions. It is located on SUS+ physical infrastructure hosted in the Crown Hosting datacentres • GPET-E: The **GP Extraction Tool-Extractor** are the GP System Supplier solutions used to	Breach of security resulting in unauthorised disclosure of personal data. (Risk 1,5) Risk that data are repurposed once the COVID-19 Direction expires and it is not transparent to patients (Risk 11)

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			extract the data from the clinical system on receipt of the request from the GP Data Collector. The resulting data files are sent to the NHS Digital DPS MESH mailbox • MESH: **Message Exchange for Social Care and Health** is the secure transport mechanism used to transport the data from the GPET-Es to NHS Digital. Data files are stored on MESH in accordance with MESH's 30-day retention policy, and are then deleted from MESH. DPS is the platform where the data will be processed and stored. NHS Digital uses Amazon Web Services (AWS) to host the data located within the UK, consequently AWS is a data processor for all data stored on DPS and NHS Digital has GDPR Article 28(3) compliant contracts in place with AWS. See Appendix C for more detail on security. The overall data flow is detailed in Section 2 Part C-4. Data will be retained in accordance with the records management policy[4] of NHS Digital. The ongoing collection and analysis of the data will continue until the expiry of the COVID-19 Direction. This is currently 31 March 2022 but will be reviewed in September 2020 and every 6 months thereafter. NHS Digital will retain the data collected for the following purposes:	

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

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			• To make it available to approved organisations who continue to require it for COVID-19 planning and research purposes and who have a legal basis to process it • For internal audit and legal record keeping purposes in relation to the data NHS Digital itself has analysed under the COVID-19 Direction and in relation to the data disseminated to third parties. Further information on the storage and transmission arrangements for Collected data are set out in **Part B-3 (Description of the Data Flows)**.	
4	Nature of Processing Compliance with Integrity and Confidentiality Principle	Who can access the information collected during this project? (including any other third parties etc.)?	**External Parties** All requests to access the information will be made via NHSX who will be responsible for a single triage service to request access to health and care data in order to support the COVID-19 response that will triage and prioritise applications that are applicable to this data set. Those applications will then be passed to NHS Digital who will be responsible for assessing and fulfilling the applications; these applications will only be successful if they pass the appropriate ethical, legal and Information Governance requirements. This is to ensure that data is only shared where it is secure, lawful and appropriate to do so. NHS Digital will do this through the Data Access Request Service (DARS) with advice on requests for record level data from this	Access arrangements are broader than necessary and result in increased risk that the Collected Data are used for unauthorised purposes, (i.e. not the COVID-19 purposes). (Risk 6, 14) Risk that Collected Data are used for purposes that are not expected but are legitimate COVID-19 purposes. (Risk 13) Risk that Collected Data are subject to unauthorised access, or are lost, damaged or stolen (Risk 5)

			collection from the Independent Group Advising on the Release of Data (IGARD). Data recipients will need to demonstrate they have a lawful basis to process the data for COVID-19 purposes and NHS Digital will need to have a lawful basis to share the data with them for that purpose as determined through the DARS process. Further detail on the methods for data access are described in Part C, Part 4, Stage 5. The Data Access Environment (DAE) is the default method of dissemination for GDPPR data to approved customers. The DAE is a secure area within the DPS platform which presents the data stored within DPS using Views of the data which have been developed and made available as security-controlled data assets. It is a secure way for users to remotely access better linked information, while ensuring that the right person, with the right permissions gets the right data, in accordance with their Data Sharing Agreement (DSA). DAE provides a single access environment for internal and external users to access this data and supports a number of presentation tools. Internal usage is described below in Internal Parties. By default, users cannot download the results of queries performed on the data from DAE. However, there are cases where this is necessary in which case the user is granted specific permission to download data as part of the DARS process. More information is available in **Section 2 Part C Question 4 Stage 5** and **Appendix C**.	

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			DAE protects data using the following methods: • DAE enforces multi-factor authentication for all user access; • DAE applies access controls on user access restricting available data to that which has been authorised under a DARS agreement or in line with the process governing access by analysts; • DAE uses virtual desktops, providing users with a view of the desktop – this means all processing is within DPS and that as part of normal processing data is never downloaded to a user's PC; • A user can only download data from DAE where this has been specifically authorised via the DARS process. By default, downloads are not permitted. Access to the download tool is controlled via multi-factor authentication of the user, and the user can only access results to their queries (see Section 2 Part C Question 4 Stage 5 DAE); • Each user is linked to one or more Data Views, but can only access one Data View at a time; • Each Data View uses a different set of pseudonyms for identifiers, thus preventing	

			unauthorised linkage of data across different Data Views. In certain scenarios, users may look to receive a data file rather than access the data on-line through DAE. Such requirements must be approved through their DARS Data Sharing Agreement and will be serviced by NHS Digital using DPS to produce extract files for recipients of the data. The data is protected using the following methods: • The data files will be sent to the recipient using a secure mechanism such as MESH or SEFT; • The data files will only contain data that has been approved through the DARS process; • It may be necessary to deliver some disseminations in multiple data files, in this case each file will use the same set of pseudonyms; • Each Data View uses a different set of pseudonyms for identifiers, thus preventing unauthorised linkage of data across different Data Views. **Internal Parties** Only approved analysts in NHS Digital will have access to the Collected Data held in DPS for appropriate and necessary data management, preparation and analysis for COVID-19 related purposes.	

			This includes the NHS Digital analytics team carrying out approved internal analysis for COVID-19 purposes. Permission to carry out any internal analysis is given by the Information Asset Owner and the Information Governance Team who will assure that there is a legal basis for the analysis to be carried out. It also includes the creation of data derivations items (e.g. creating Year of Birth from Date of Birth) to provide non-identifiable data items for customers and performing the relevant linkage for customers as approved by DARS. Access to the data is strictly limited and subject to authorisation by the Information Asset Owner through NHS Digital's own Clear Data Access internal approval process. This process ensures that authorisation is only given to an individual for a time-limited period, where the access to the data is justified, and an appropriate legal basis for such processing is in place. Where NHS Digital carries out analysis for statistical purpose and publishes anonymous statistical data, the resultant aggregated information will have any small numbers suppressed in accordance with the Code of Practice for Statistics, and is thus anonymous. Publication of the data must either be agreed by the Secretary of State (or an officer acting on his behalf) or NHS Digital must reasonably believe it to be in the public interest to publish the data, following consultation with relevant parties, for example NHSX, Public Health England and professional bodies (this would include the BMA, the RCGP and the JGIPTC).	

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			All new statistical releases of data are reviewed by NHS Digital's Disclosure Control Panel made up of senior statisticians as well as the Chief Statistician to ensure data is not identifiable (e.g. by suppressing small numbers, aggregating data to large geographical areas or using other methods of disclosure control).	
5	Nature of Processing Compliance with Accountability Principle	Identify all external third parties will have some involvement in the project (i.e. they may have access to, store or otherwise process the personal data) Explain what their involvement / role will be in the project.	The following: • The GP Practice system supplier will perform the data extract using GPES. They are data processors for GPs and will be instructed by GPs to carry out the extract through GPs accepting the invitation to participate in the Calculating Quality Reporting Service (CQRS). • Data customers, with a legal basis to access the data approved by DARS (with IGARD oversight). These are likely to include research bodies, CCGs and NHS bodies such as PHE and NHSE and NHSI. Data processors acting on their behalf will be identified as part of any application process. • Amazon Web Services provides the IT Platform Infrastructure and is therefore our data processor for data stored in DPS and the DAE. They have been appointed as a data processor under the	Risk that there are insufficient data protection controls around third parties who have access to the Collected Data. (Risk 5)

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			terms of the Softcat Cloud Billing contract which contains Article 28(3) compliant terms. • Contractors, seconded staff or those on honorary contracts who are processing data for NHS Digital as part of their work. Their need for access to data will be considered as part of the Clear Data Access Process.	
6	Nature of Processing Compliance with Accountability Principle	Please identify any internal stakeholders this project has been discussed with Explain to which extent they are involved In the Project.	NHS Digital COVID -19 Silver and Gold Commands for COVID-19 programme – the Commands have provided sponsorship and executive level oversight up to and including the Chief Executive. NHS Digital Data Protection Officer - the DPO has been involved in this project and has embedded a member of the DPO team on the project to report back on progress. Key documents have also been reviewed by the DPO including the legal basis for processing, compliance with the principles and transparency material. NHS Digital Data Access Request Service (DARS) – responsible for enabling appropriate and secure access to GDPPR data; has provided guidance to and worked with the GDPPR team to ensure that appropriate actions were taken to achieve this. NHS Digital Data Processing Services (DPS) – works closely with GDPPR team to develop functionality in DPS to ensure that the data will be landed, processed,	Due to extent of involvement of key control areas within NHS Digital as identified here there is not considered to be a material residual risk arising from insufficient internal stakeholder involvement.

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			accessed and linked in line with information governance, business processes and technical standards and requirements to meet the aims of GDPPR. NHS Digital Data Services Alliance – provides development teams for GP Data Collector (part of GPES), DPS and DAE and set up access to the various DPS/DAE tools to users of the DPS platform. NHS Digital Caldicott Guardian – provides support, advice and guidance on confidentiality and ethical issues for GDPPR and liaised directly with BMA and RCGP representatives. GPES Information Asset Owner (IAO) – the IAO is accountable for the management of the data asset within NHS Digital and has provided leadership and taken decisions regarding all aspects of this project, particularly those concerning access, risk and documentation. NHS Digital Live Services – the team provides service management for the GP IT System Suppliers delivery of the GDPPR data, provides service management of the DPS platform and manages access requests to DAE. NHS Digital Cyber Security - this team ensures that patient data is safely and securely stored and has been consulted for assurance as to the security of DPS and DAE.	

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			NHS Digital IG Team - this team has provided support, guidance, leadership and liaison with external stakeholders on all matters relating to information governance, compliance and transparency, up to and including Executive Director level.	
7	Nature of Processing Compliance with Storage Limitation Principle	Where geographically will the personal data be stored, or otherwise processed by any of the identified third parties? If the processing will involve processing outside of the UK, where will the processing take place and what safeguards apply under Articles 45-49 GDPR?	The UK. The geographical location of the data once transferred to NHS Digital data is within the UK jurisdiction in the DPS Platform, which is hosted in the AWS cloud within the UK. All NHS Digital staff and contractors accessing the DPS will be doing so from within the UK. For non-NHS Digital use, whether for data files or on-line access within the DAE, the DARS application specifically states locations for storage and processing, and the Territory of Use. This is captured within the Data Sharing Agreement, which restricts use to these addresses and the territory. The application process includes assessing the recipient's legal basis for processing within these territories, and in particular where such the territory of use is outside the UK	Risk that Collected Data are transferred outside of the UK in breach of GDPR restricted transfer requirements due to ineffective monitoring of processing activity, including of disseminated data. (Risk 12)
8	Nature of Processing Compliance with Storage	When will personal data be deleted, and how?	The extraction will continue until the expiry of the COVID-19 Direction. This is currently 31 March 2022 but will be reviewed in September 2020 and every six months thereafter. It may be extended by the Secretary of State if data is still required to support the response to COVID-19 beyond this date. Any	Risk that personal data is held for longer than necessary for the purposes set out in the COVID-19 Direction or other lawful authority considered in this DPIA. (Risk 9)

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
	Limitation Principle		extension will be published here. In line with NHS Digital records management policy, data will be kept for for 8 years after last use for legal reasons to enable NHS Digital to exercise and defend its legal rights in relation to the data or any actions taken by NHS Digital eg. analysis and dissemination (legal purposes). Data which has been disseminated will be held by the Data Recipient for the purposes and duration permitted under the relevant Data Sharing Agreement. This will be assessed as part of the DARS process.	Risk that data are repurposed once the COVID-19 Direction expires. (Risk 11)
9	Accuracy Principle	How will the personal data be amended and kept up to date?	There will be fortnightly snapshots in time extractions via GPES until the expiry of the COVID-19 Direction. These snapshots will contain the full GDPPR data set for any patient record that has had a SNOMED code added to their record since the previous extract, has newly registered at a GP practice or has died since the last extract, as outlined in the 'background' section of the DPIA above. These patient records will be merged into the data set stored in DPS as part of the data ingestion pipeline.	Risk that the Collected Data is not up to date at the time of dissemination due to the time lags between the scheduled extractions and processing of the data into DPS. (Risk 10).
10	Scope of Processing	How many individuals are likely to be affected by the project?	The GPES data extraction will include all patients in England who: • are currently registered with a general practice or who were registered with a general practice and	Risk that more data than is necessary for NHS Digital's purposes will be shared by GP Practice. (Risk 3). Risk that access rights are given to the data to those who

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
		Where are they located?	have a date of death on or after 1 November 2019; and • have SNOMED codes in their patient records matching any of the relevant codes. It is estimated that approximately 54 million patients will be included in the cohort of Collected Data.	do not have an authorised COVID-19 purpose. (Risk 14)
11	Nature of Processing Compliance with Integrity and Confidentiality Principle	What security measures will you have in place to protect the personal data being processed?	**Security regarding transmission of data to and storage by NHS Digital** Data will be extracted by GP systems suppliers using the GPES solution outlined in the answer to question 3 above, which is an approved and established secure mechanism for extracting and delivering data. Once GPES has collected the data, it will pass it to a DPS AWS S3 bucket. This bucket enables security cleared role-specific access only. The data is then processed into a secure database, where it is kept separate from all other data sets stored within DPS. Backups of the data will be supported by the DPS backup and recovery processes. DPS has a System Level Security Policy (SLSP) in place. Further detail is included in Section 2, Part C-4, and Appendix C.. AWS Cloud data centres 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	Breach of security in the transmission of the data to NHS Digital and storage and other processing of the Collected Data which results in the unauthorised disclosure of personal data. (Risk 5) Risk that data will be disseminated to organisations that do not meet the required security standard expected by NHS Digital. (Risk 7) Risks and mitigations relating to security measures following transmission to NHS Digital are considered in Part E of this document.

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			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/. When data is stored persistently, industry standard AES-256 encryption is used with a combination of locally managed keys and AWS supplied keys. 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) Most services use ELBSecurityPolicy-TLS-1-1-2017-01 with some newer ones adopting ELBSecurityPolicy-TLS-1-2-Ext-2018-06. These policies enforce AES-256 at minimum. GPES (GPDC) is hosted on the SUS+ physical infrastructure which is hosted in the Crown Hosting datacentres. The hosting environment is ISO27001 compliant (https://www.gov.uk/guidance/the-crown-hosting-data-centres-framework-on-the-digital-marketplace). No data is stored within GPDC, which only acts to initiate requests for data extracts. SUS+ has a System Level Security Policy (SLSP) in place.	

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			The AWS infrastructure on which DPS is deployed is continuously monitored by the Care Security Operations Centre using AWS GuardDuty rules. Unexpected and novel changes to infrastructure are identified and reported to the Security Officer. **Security regarding dissemination of data by NHS Digital** Data applicants will need to demonstrate through the DARS[5] process that they have adequate security measures in place to protect the data where it is to be disseminated to them. Such security requirements are set out as part of the DARS application process and are documented on-line within the DARS section of the NHS Digital website. The Data Access Environment (DAE) will be used for all dissemination where possible, as this is the default method of dissemination for GDPPR data. Both access to data through DAE, or for data files sent to customers, will be subject to the same level of consideration through the DARS process. Where data is required to be pseudonymised this will be performed within DPS using the national pseudonymisation tool (Privitar). DPS/DAE has a System Level Security Policy (SLSP) in place.	

[5] Information about DARS and the application process is available at https://digital.nhs.uk/services/data-access-request-service-dars

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			Further detail on DPS/DAE security is included in Section 2, Part C Part 4 and in Appendix C.	
12	The Scope of Processing Compliance with Storage Limitation Principle	How long will the processing be taking place (i.e. for the duration of the project (provide number of weeks/months/years) and whether the processing will continue beyond the end of the project)?	This data will be extracted as a snapshot in time extract on the initial collection. A subsequent fortnightly snapshot extraction will then continue until the expiry of the COVID-19 Direction. This is currently 31 March 2022 but will be reviewed in September 2020 and every six months thereafter but may be extended by the Secretary of State if data is still required to support the response to COVID-19 beyond this date. Any extension will be published here. Data will be retained in accordance with the records management policy of NHS Digital and the Records Management Code of Practice for Health and Social Care 2016[6]. NHS Digital will retain the data collected for 8 years from the date of last collection for the following purposes: • To make it available to approved organisations who continue to require it for COVID-19 planning and research purposes and who have a legal basis to process it • For internal audit and legal record keeping purposes in relation to the data NHS Digital itself has collected and analysed under the	The personal data is held for longer than necessary for the purposes set out in the COVID-19 Direction or other lawful authority considered in this DPIA. (Risk 9)

[6] https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/codes-of-practice-for-handling-information-in-health-and-care/records-management-code-of-practice-for-health-and-social-care-2016#

No	Compliance Requirement	Question	Answer	Risks: To be mitigate as set out in Part E
			COVID-19 Direction and in relation to the data disseminated to third parties. Further information on the storage and transmission arrangements for Collected data are set out in **Part B-3 (Description of the Data Flows)**.	
13	The Nature of Processing Compliance with Integrity and Confidentiality and Data Minimisation Principle.	Will data be anonymised, Pseudonymised, aggregated in non-personal information, etc? If so, please explain whether this is only externally or also internally within the organisation.	Data will be pseudonymised by use of the national pseudonymisation tool Privitar as explained above. Data disseminated directly could be in anonymised, pseudonymised or identifiable form, according to what is necessary to meet the approved purpose and what legal basis the requestor has to process the data, which is assessed as part of the DARS process. Further detail is included in Section 2, Part C Part 4.	Ineffective or failed pseudonymisation at dissemination and/or re-identification of data outside the DARS approved use and legal basis of a data recipient. (Risk 1) the COVID-19

4. Describe the personal data flows

Table 1 below and summarise the end to end (GPSS) to the data being made available to end-users by NHS Digital via the Data Processing Service (DPS)[7]. The following sections describe it in more detail.

Table 1: GPES Data for Pandemic Planning and Research Data Flow

Category and Description of Processing	Entity involved	Types of Personal Data Processed	What are the Data Processing and security arrangements
Stage 1: GP IT System Suppliers extract data to NHS Digital	Data Controller: GP Practice Data Processor: GPSS	Identifiable patient data	Data are held in GP patient record systems which meet national standards for security under relevant GP IT System contracts. Identifiable patient data will be extracted via the General Practice Extraction Tool within GPES. [Exceptions process] (a) In the event that the extract process fails, the GP system supplier will detect this and stop the process, thus preventing the data being transferred to NHS Digital.
Stage 2: GP IT System Suppliers provide data to NHS Digital	Data Controller: GP Practice & NHS Digital Data Processor: GPSS	Identifiable patient data	Data will be transferred to NHS Digital through an approved secure method of the Message Exchange for Social Care and Health (MESH). See Stage 2 description below for further details) [Exceptions process] (a) In the event that the transfer of data to NHS Digital fails, then the transfer will be repeated

[7] See Appendix C for a description of DPS

Category and Description of Processing	Entity involved	Types of Personal Data Processed	What are the Data Processing and security arrangements
Stage 3: NHS Digital receives data into the Data Processing service	Data Controller: NHS Digital	Identifiable patient data	Data are processed in DPS to verify completeness and accuracy of data [Exceptions process] (a) In the event that completeness and validation checks fail, then the cause will be investigated and, when resolved, the process will be restarted from stage 1 with the regeneration of the data extract.
Stage 4: NHS Digital de-identifies the data	Data Controller: NHS Digital	Identifiable patient data	De-identification rules are applied to the data, and the data saved in the secure environment [Exceptions process] a) In the event that de-identification fails, then the cause will be investigated and, when resolved, the process will be restarted from Stage 3.
Stage 5: Data is made available to end users	Data Controller: NHS Digital	Pseudonymised patient data or identifiable patient data	Users can view and analyse the pseudonymised data or clear (identifiable) data, as per their authorisation via the DARS process. [Exceptions process] (a) In the event that a user attempts to view data that they have not been authorised to access via the DARS process, then the attempt is blocked, and no data is made available.

Stage 1: GP IT System Suppliers extract data to NHS Digital

GPES is used to generate the identifiable patient data extract from participating GP practices.

Stage 2: GP IT System Suppliers provide data to NHS Digital

The patient data is transferred from the GP System Supplier (GPSS) to the NHS Digital Data Processing Service (DPS) using the Message Exchange for Social Care and Health (MESH) service for secure large file transfers. The same mechanism will be used to send submission acknowledgements (stages 3 and 4) back to the GP System Supplier when DPS has finished processing the data file.

MESH is a service provided by NHS Digital and is the main secure large file transfer service used across health and social care organisations for clinical and other data.

The steps to transfer the data file are:

1. The GPSS submits the GDPRR extract file to MESH service via their MESH Client
2. The MESH service routes the data file to DPS
3. DPS downloads the data file using the DPS MESH client to the DPS Landing Zone file store.

The steps to transfer the submission acknowledgement file from DPS are:

1. DPS submits the acknowledgement file to the MESH service via the DPS MESH Client
2. The MESH service routes the acknowledgement file to the GPSS
3. The GPSS downloads the acknowledgement file using their MESH Client.

Security:

- MESH authenticates all client connections using TLS Mutual Authentication
- MESH ensures that the extract file is encrypted while it is being transferred from the GPSS to DPS – all connections to and from the MESH service use https.

Stage 3 and 4: DPS processing of data

DPS processes submitted data and prepares it for authorised use. The processing steps are:

1. DPS takes the extract file from the Landing Zone file store and applies validation and data quality (DQ) checks.

2. DPS calls the De-Identification Service to tokenise identifiers to the DPS internal pseudonyms
3. DPS calls the De-Identification service to re-tokenise the internal pseudonyms to those that will be seen by users, for example authorised NHS Digital analysts
4. DPS updates the GDPRR asset using the submitted file:
    a. For new patients, the data is added to the GDPRR asset
    b. For updated patient data, the existing DPS held data is updated with the changes.
5. DPS updates the pseudonym mapping tables with internal to recipient pseudonym mappings
6. DPS generates a submission validation report for return to the GPSS (see stage 2).

Security:

- DPS internal pseudonyms are never exposed to users
- All file stores used by DPS are encrypted (see Appendix C.2).

Stage 5: Data is made available to end users

All requests for data held by NHS Digital, including GDPRR data, for aggregate, identifiable and pseudonymised will be accessed through the Data Access Environment (DAE) within DPS, excepting those cases where DAE cannot support the requirements of that customer at that time and/or where data must be disseminated to the customer based on specific need.

All access to GDPRR data must be approved via the DARS process and a Data Sharing Agreement is in place.

Each DARS approved request will use a different set (Domain) of pseudonyms, this reduces the risk of unauthorised linking of data between organisations. Figure 1 shows different data views using different sets of pseudonyms and the dissemination of data either via DAE or via data files. These mechanisms are described below.

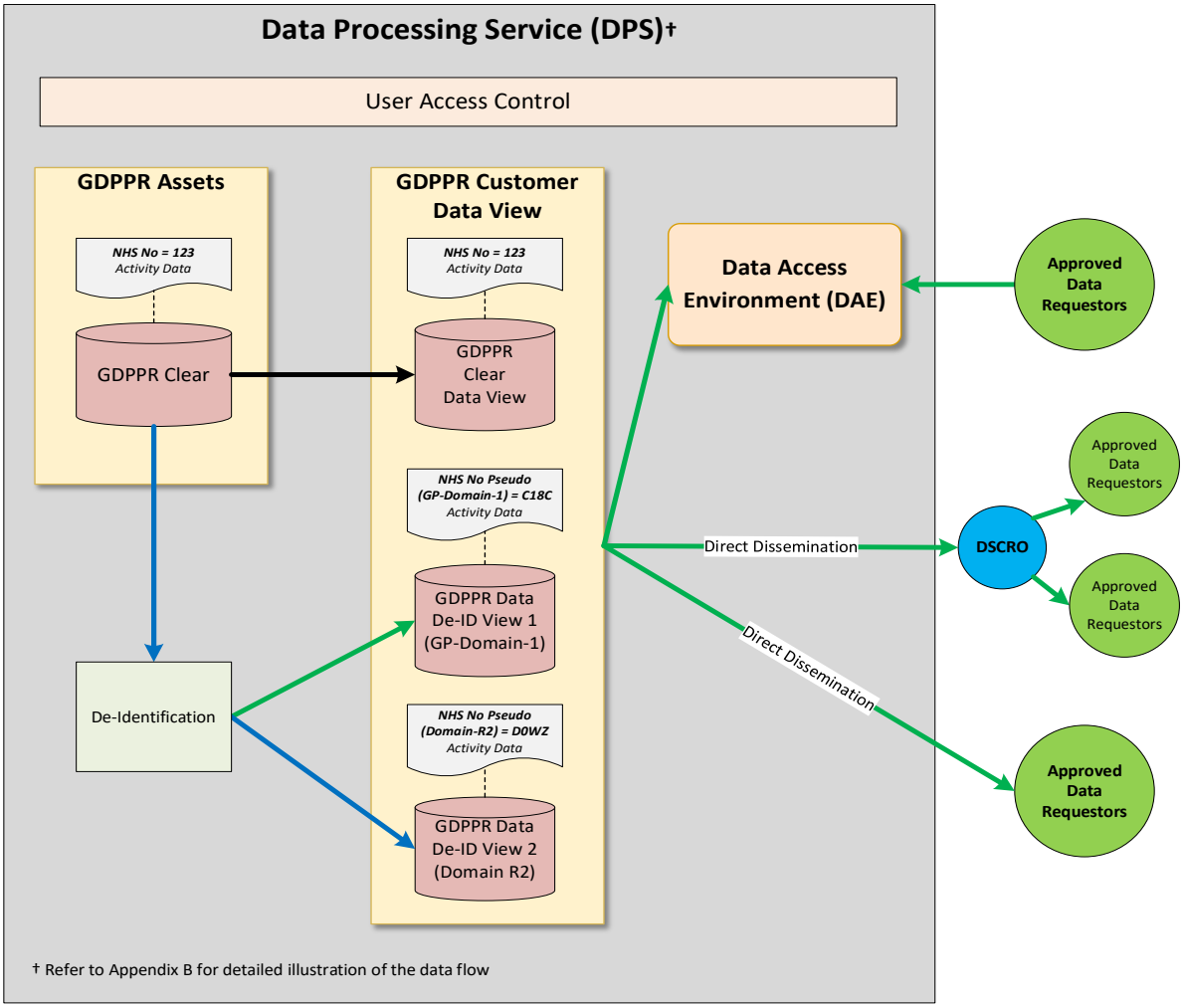

Figure 1: Data Views and Disseminations

Where users access the data within the DAE and where a user is not entitled to see Personal Identifiable Data, the linkage will be done by joining datasets with the appropriate domain pseudonym. Domain pseudonyms will only be generated for the datasets covered by their associated Data Sharing Agreement.

Where identifiable GDPPR data is sent to data requesters and can be linked to other datasets that the requester holds, under the permissions of the Data Sharing Agreement, the requester should link datasets using NHS Number and Date of Birth. When deidentified, data requesters should use the domain pseudonym to link datasets.

Data View Security:

- Each approved DARS request is represented by a different data view. The data view is a control which limits the data available to that which has been authorised.

- Each data view will use a different set of pseudonyms (based on the approved DARS request)
- Each user is linked to one or more data views and can only access one data view at a time.

Data minimisation

Personal data will be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed. This means the data requester must justify the amount they can view by the purpose asserted within the application. Additional justification is required where necessary. More information can be found at https://digital.nhs.uk/services/data-access-request-service-dars/dars-guidance/data-minimisation.

Data Access Environment (DAE)

DAE is a single access environment for NHS Digital and external users to access this data which supports a number of presentation tools. By default, users cannot download the results of queries from DAE. However, there are cases, typically involving cohort management, where this is necessary in which case the user is granted specific permission to download data.

Security:

- DAE enforces multi-factor authentication for all user access
- DAE applies access controls on user access restricting available data to that which has been authorised via a Data Sharing Agreement (DSA) which has been approved through the Data Access Request Service (DARS)
- DAE uses virtual desktops, providing users with a view of the desktop – this means all processing is within DPS and that, as part of normal use, data is never downloaded to a user's PC.
- A user can only download data from DAE where this has been specifically authorised as part of the DARS and IGARD process and is subject to multi-factor authentication.

Data Dissemination

DPS produces a data file based on the recipient's Data Sharing Agreement. This will use a set of pseudonyms unique to the recipient. The data file is transferred to the recipient using a secure mechanism.

Security:

- The file will be sent to the recipient using a secure mechanism such as MESH

- The data file will only contain data that has been authorised via a Data Sharing Agreement (DSA) which has been approved through the Data Access Request Service (DARS)
- Each recipient will receive data with a different set of pseudonyms (based on the DSA)

PART D: NECESSITY AND PROPORTIONALITY

1. What is your lawful basis for processing the personal data?

NHS Digital Statutory Authority for Collection and Analysis of Data		
Are NHSD a Controller or Processor of this Data?	Controller, jointly by law with the Secretary of State for Health and Social Care in relation to determining the overarching COVID-19 purposes for the collection, analysis and dissemination of the data obtained under the COVID-19 Public Health Direction 2020 (see below)	
Statutory Authority for NHSD to hold the data requested.	Direction: Covid-19 Public Health Directions 2020. The purpose of collecting this data is for planning and research COVID-19 Purposes and the data collected is a COVID-19 Public Health Information System. The analysis of this data has been determined by NHS Digital to be necessary as part of the COVID-19 co-ordinated response, particularly with reference to the request from BMA and RCGP to help relieve the burden on GPs and the need for the data for vital planning and research. See DPN: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/gpes-data-for-pandemic-planning-and-research	
Compliance with Common Law Duty of Confidentiality by NHSD and Providers on Collection and Analysis		
COPI? **COPI Notice?** **Other Legal Obligation?** **Consent?**	**NHS Digital** Legal Obligation to collect data– Above COVID-19 Public Health Direction under S254 of 2012 Act	**Provider – GP Practice** Legal Obligation to provide data- S259 of Health and Social Care Act 2012. See Data Provision Notice.
GDPR Compliance for Collection and Analysis		
Article 6	**NHS Digital** Article 6(1)(c) – Legal Obligation – by virtue of COVID-19 Public Health Direction under S254 of 2012 Act 6(1)(e) – public task – COVID-19 Public Health Direction 2020 in relation to any analysis determined by NHS Digital which is not requested by DHSC/NHSX.	**Provider – GP Practice** Article 6(1)(c) – Legal Obligation to provide data – by virtue of S259 of Health and Social Care Act 2012. See Data Provision Notice.

Article 9	**NHS Digital** Article 9(2)(g) – substantial public interest, plus Part 2 Sched 1 DPA18, para 6 statutory and governmental purpose re COVID-19 Public Health Direction. NHS Digital has in place and Appropriate Policy Document in relation to processing on this basis as required by paragraph 5 of Part 2 and Part 4 of Schedule 1 of DPA18. Collection and analysis of this data by NHS Digital is in the substantial public interest due to: • the reduction in burden and cost on GP Practices and GP System Suppliers of multiple requests for GP data under COPI Notices • the single point of access process managed by DARS and overseen by IGARD of providing access to the data which will ensure it is appropriate, secure and used for appropriate purposes with a legal basis and data sharing agreement regulating use. It will increase transparency through publication of dissemination details in the NHS Digital Data release register • the need for GP data to be shared to carry our vital planning and commissioning of services to manage the COVID-19 outbreak and to carry our vital COVID-19	**Provider – GP Practice** • Article 9(2)(g) – substantial public interest, plus Part 2 Sched 1 DPA18, para 6 statutory and governmental purpose re S259 of the Health and Social Care Act 2012 and Data Provision Notice to provide data. • Article 9(2)(h) – necessary for the purposes of the management of health and social care systems, plus paragraph 2 of Part 1 of Schedule 1 of the DPA 18 (Healthcare Purposes) • Article 9(2)(i) – necessary for reasons of public interest in the area of public health, plus paragraph 3 of Part 1 of Schedule 1 of DPA 18 Public Health Purposes) • Article 9(2)(j) – scientific research or statistical purposes and paragraph 4 of Schedule 1 (research and statistical purposes). Provision of this data to NHS Digital is in the substantial public interest due to: • the reduction in burden and cost on GP Practices and GP System Suppliers of multiple requests for GP data under COPI Notices • the single point of access process managed by NHS Digital through DARS with

	research into the cause, effects, treatments and vaccines for COVID-19 as part of the response and management of the COVID-19 outbreak, protection of public health and promotion of health.	oversight by IGARD of providing access to the data which will ensure it is appropriate, secure and used for appropriate purposes with a legal basis and data sharing agreement regulating use. It will increase transparency through publication of dissemination details in the NHS Digital Data release register • the need for GP data to be shared to carry our vital planning and commissioning of services to manage the COVID-19 outbreak and to carry our vital COVID-19 research into the cause, effects, treatments and vaccines for COVID-19 as part of the response and management of the COVID-19 outbreak, protection of public health and promotion of health.
Transparency Notice Article 13 and 14	**NHS Digital** NHS Digital has drafted a specific GPES Data for Pandemic Planning and Research Transparency Notice here	**Provider** NHS Digital has drafted a template General Practice Transparency Notice for GPs to use. This is referred to in the DPN and is available here
NHS Digital Statutory Authority to Disseminate the Data and Recipient Statutory Authority to receive the Data		
COPI Notice? COPI Regs?	Case by Case Assessment but likely to be: • COPI • NHS Digital COPI Notice	Case by Case Assessment but likely to be: • COPI • COPI Notice (General or NHSE/I)

Section 261 Of the Health and Social Care Act 2012?	NHS Digital only Case by Case Assessment but likely to be: • Dissemination under COVID-19 Directions • Section 261(1) • Section 261(4) • Section 261(5)	
Other?	Case by Case Assessment E.g. Notice requiring disclosure under Section 45C of the Statistics and Registration Services Act by the Statistics Board	Case by Case Assessment E.g. Notice requiring disclosure under Section 45C of the Statistics and Registration Services Act to the Statistics Board
Compliance with Common Law on Dissemination by NHSD and Recipient		
Are we relying on COPI? COPI Notice? Section 251 Approval?	**NHS Digital** Case by Case Assessment by DARS but likely to be: • COPI • NHS Digital COPI Notice • Section 251 Approval • Patient consent	**Recipient** Case by Case Assessment by DARS but likely to be: • COPI • General or NHSE/I COPI Notice • Section 251 Approval • Patient consent
GDPR Compliance for NHSD dissemination by NHSD, use by recipient		
Article 6	**NHS Digital** Case by Case Assessment by DARS which will depend on the organisation we are sharing the data with and their purposes for using the data. This will include: • Article 6(1)(c) – legal obligation, for example where the NHS Digital COPI Notice applies • Article 6(1)(d) – vital interests, for example where it is necessary to protect patients' vital interests • Article 6(1)(e) – public task, for example where we are sharing data with	**Recipient** Case by Case Assessment by DARS which will depend on the organisation we are sharing data with and their purposes for using the data. This will include: • Article 6(1)(c) – legal obligation, for example where the General COPI Notice applies • Article 6(1)(d) – vital interests, for example where it is necessary to protect patients' vital interests • Article 6(1)(e) – public task, for example where

	another public authority for the purposes of them exercising their statutory or governmental functions • Article 6(1)(f) – legitimate interests, for example where we are sharing information with a research organisation to carry out coronavirus research	we are sharing data with another public authority for the purposes of them exercising their statutory or governmental functions • Article 6(1)(f) – legitimate interests, for example where we are sharing information with a research organisation to carry out coronavirus research
Article 9 and relevant DPA Schedule 1 Condition	**NHS Digital** Case by Case Assessment by DARS but likely to be: • Article 9(2)(g) – substantial public interest, for the purposes of NHS Digital exercising its statutory functions or for other organisations exercising their governmental or statutory functions, and paragraph 6 of Schedule 1 DPA 18 (statutory purposes) • Article 9(2)(h) – health or social care purposes and paragraph 2 of Schedule 1 of DPA18 (healthcare purposes) • Article 9(2)(i) – public health purposes and paragraph 3 of Schedule 1 of DPA18 (public health purposes) • Article 9(2)(j) – scientific research or statistical purposes and paragraph	**Recipient** Case by Case Assessment by DARS but likely to be: • Article 9(2)(g) – substantial public interest, for the purposes of the data recipient exercising its governmental or statutory functions, and paragraph 6 of Schedule 1 DPA 18 (statutory and governmental purposes) • Article 9(2)(h) – health or social care purposes and paragraph 2 of Schedule 1 of DPA18 (healthcare purposes) • Article 9(2)(i) – public health purposes and paragraph 3 of Schedule 1 of DPA18 (public health purposes) • Article 9(2)(j) – scientific research or statistical purposes and paragraph 4 of Schedule 1 (research and statistical purposes).

	4 of Schedule 1 (research and statistical purposes) It is likely that the disclosure to the recipient will be in the substantial public interest due to the need to carry our vital planning and commissioning of services to manage the COVID-19 outbreak or to carry our vital COVID-19 research into the cause, effects, treatments and vaccines	It is likely that the disclosure to the recipient will be in the substantial public interest due to the need to carry our vital planning and commissioning of services to manage the COVID-19 outbreak or to carry our vital COVID-19 research into the cause, effects, treatments and vaccines

2. Is it necessary to collect all data items to achieve the purpose of the Project?

Below is a table outlining the justification for necessity to collect the data items in this collection.

Data Items/Categories	Current, historic and any time restriction	Description and Justify why the item is being collected
Personal Data		
NHS Number	Current data only	NHS Number will be provided by the GP system supplier. A pseudonymised version will be generated by NHS Digital to be used as de-identified alternative where possible (further detail is included in Section 2, Part C-4). To enable planning and research in relation to COVID-19 and linkage with other data sets to enable broader research. Where the NHS Number field is blank, DPS may use other demographic detail collected (e.g.: surname as detailed below) in combination in an attempt to find the related NHS Number if it is decided that there is a problem with missing NHS Numbers.
Surname	Current data only	Surname will be provided by the GP system supplier. If the NHS Number field is empty then

Data Items/Categories	Current, historic and any time restriction	Description and Justify why the item is being collected
		Surname is one of the fields that can be used by NHS Digital to match the patient data to the correct NHS Number. There is the potential that this data item may be requested for planning and research in relation to COVID-19 for an identifiable cohort of patients e.g. patients who have consented to take part in a specific research study.
Forename	Current data only	Forename will be provided by the GP system supplier. If the NHS Number field is empty then Forename is one of the fields that can be used by NHS Digital to match the patient data to its NHS Number. There is the potential that this data item may be requested for planning and research in relation to COVID-19 for an identifiable cohort of patients e.g. patients who have consented to take part in a specific research study.
Address	Current data only	Address will be provided by the GP system supplier. If the NHS Number field is empty, then Address is one of the fields that can be used by NHS Digital to match the patient data to its NHS Number. There is the potential that this data item may be requested for planning and research in relation to COVID-19 for an identifiable cohort of patients e.g. patients who have consented to take part in a specific research study. Address or partial address also enables stratification by geographic locations where

Data Items/Categories	Current, historic and any time restriction	Description and Justify why the item is being collected
		postcode is not sufficiently granular, and thus research into specific outbreaks of disease, incidence rates, association with deprivation and general epidemiological analysis.
Postcode	Current data only	Postcode will be provided by the GP system supplier. Lower Layer Super Output (LSOA) will be derived from this value by NHS Digital to be used as a minimised alternative where possible. LSOA is also used as a deprivation index. To enable planning and research in relation to COVID-19 e.g. postcode enables stratification by geographic locations and thus research into specific outbreaks of disease, incidence rates, association with deprivation and general epidemiological analysis.
Date of Birth	Current data only	Date of Birth will be provided by the GP system supplier. Year of Birth will be derived from this value by NHS Digital to be used as a minimised alternative where possible. To enable planning and research in relation to COVID-19 e.g. DOB enables stratification by age, and general epidemiological analysis.
Date of Death	Current data only	Date of Death will be provided by the GP system supplier. Year of Death will also be derived from this value by NHS Digital to be used as a minimised alternative where possible, as per the ICO Anonymisation Code which states that clear data could potentially be used by a 'motivated intruder'

Data Items/Categories	Current, historic and any time restriction	Description and Justify why the item is being collected
		alongside other available data to reidentify someone. To enable planning and research in relation to COVID-19 e.g. Date of Death enables establishment of cohorts of fatalities and general epidemiological analysis.
Sex	Current data only	Sex will be provided by the GP system supplier. To enable planning and research in relation to COVID-19 e.g. Sex enables stratification by this item and thus research into incidence rates, association with deprivation and general epidemiological analysis and potential inequalities (e.g. trends in disease prevalence for males).
Physical Description	Current and the last 2 years of values relating to Physical Description as specified for relevant SNOMED codes in the business rules	The Patient Record may contain clinical codes that could relate to the individual's physical description such as 'obese' where these occur within specified SNOMED codes. To enable planning and research in relation to COVID-19 e.g. research into factors associated with prevalence and general epidemiological analysis.
Special Category Data		
Physical / Mental Health or Condition	Current and the last 2 years of values relating to Physical/Health Condition as specified for relevant SNOMED codes in the business rules	The Patient Record may contain clinical codes that relate to the individual's physical and/or mental health/conditions such as 'depression' where these occur within specified SNOMED codes. To enable planning and research in relation to COVID-19 e.g.

Data Items/Categories	Current, historic and any time restriction	Description and Justify why the item is being collected
		research into factors associated with prevalence and general epidemiological analysis.
Sexual Life	Current and the last 2 years of values relating to Physical/Health Condition as specified for relevant SNOMED codes in the business rules	The Patient Record may contain clinical codes that relate to the individual's physical and/or mental health/conditions which reference a sexually transmitted infection (STI), human immunodeficiency virus (HIV), acquired immune deficiency syndrome (AIDS) or a gender related disorder, because the code occurs within relevant SNOMED codes in the business rules, e.g. these are codes relating to respiratory conditions or cancers which are relevant in the context of COVID-19 research and planning. Therefore, the extract does contain sensitive codes (171) which are listed in the SNOMED reference tables of codes relating to sensitive information. To enable planning and research in relation to COVID-19 e.g. research into factors associated with prevalence such as compromised immunity, and general epidemiological analysis.
Ethnicity	The current value on the GP record and all historic values.	Ethnicity will be provided by the GP system supplier. To enable planning and research in relation to COVID-19 e.g. Ethnicity enables stratification by this item and thus research into incidence rates and potential inequalities (e.g. trends in disease prevalence for black and minority ethnic groups) and general epidemiological analysis.

3. How have you complied with the Data Minimisation Principle?

NHS Digital has engaged with GP professional bodies, research organisations and GP IT System Suppliers to establish the minimum necessary data set to meet established and developing use cases, as set out in the DPN and Business Rules. https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/gpes-data-for-pandemic-planning-and-research

This includes:

- limitations where historic data is not required to meet uses due to its age:
    - only current values for patient demographic data are collected (exception being Ethnicity where historic values are also collected);
    - only historic values for the last two years for data relating to measurements, tests and interventions are collected;
- the creation of derived or pseudonymised versions of identifiable data items which should be used to meet the purpose of dissemination requests wherever possible to make the data less precise than the original data but sufficient for analyses, thereby mitigating the risk of patients being identified. These are:
    - NHS Number (pseudonymised from NHS Number);
    - Lower Super Output Area (derived from Postcode);
    - Year of Birth (derived from Date of Birth);
    - Year of Death (derived from date of Death);
- legally restricted codes for Gender Recognition and Human Fertilisation and Embryology will not be collected as these codes are not included in the business rules issued to GP IT System Suppliers.
- 171 sensitive codes will be collected, including codes which are listed in the 'Sexually Transmitted Disease' and the 'Gender Related' SNOMED reference tables of codes relating to sensitive information[8]. The justification for collecting these codes is set out in the table above under the Data Item/Category of 'Sexual Life'. The list of sensitive codes which occur within GDPPR is available within the guidance made available to analysts [https://digital.nhs.uk/coronavirus/gpes-data-for-pandemic-planning-and-research/guide-for-analysts-and-users-of-the-data]. Any sensitive coding may only be analysed and disseminated as part of a data release including identifiers but only where the purpose justifies that release, and there is an appropriate legal basis. Otherwise any analysis or dissemination of data which contains these codes will be aggregated, or pseudonymised with additional technical and organisational controls such that the data may be considered

[8] The reference tables of codes relating to sensitive information can be found by searching the NHSD SNOMED browser and looking at the members, or downloading them in spreadsheet form from https://isd.digital.nhs.uk/trud3/user/guest/group/0/pack/40.

anonymised in that context. All sensitive codes are separated from standard codes within DPS to enable these to be handled appropriately.The list of sensitive codes which occur within GDPPR is available within the guidance made available to analysts [https://digital.nhs.uk/coronavirus/gpes-data-for-pandemic-planning-and-research/guide-for-analysts-and-users-of-the-data]. Any sensitive coding may only be analysed and disseminated as part of a data release including identifiers but only where the purpose justifies that release, and there is an appropriate legal basis. Otherwise any analysis or dissemination of data which contains these codes will be aggregated, or pseudonymised with additional technical and organisational controls such that the data may be considered anonymised in that context. All sensitive codes are separated from standard codes within DPS to enable these to be handled appropriately.

It should be noted that the GPES tool only collects structured and clinically coded data (e.g. free text, images and documents are not collected).

4. What steps have you taken to ensure individuals are informed about the ways in which their personal data is being used for this project so as to ensure that processing is lawful, fair and transparent?

Detailed information regarding all aspects of this data collection has been made publicly available, including:

- A specific NHS Digital GDPPR Transparency Notice has been drafted which provides details regarding how, and why, NHS Digital will process and use patients' personal data in this data collection including information about the rights available to individuals to exercise. This is supplementary to our COVID-19 Response Transparency Notice and our General Transparency Notice which are all linked to the GDPPR Transparency Notice. Following consultation with patient representative groups, this Transparency Notice will be further updated.
- A specific GP Transparency Notice for this collection has also been drafted which can be used by GPs to provide transparency to their patients through publication on their website. This also provides full details including rights individuals have.
- A Data Provision Notice to GPs. This covers all aspects of this data collection, including the purpose, benefits, legal basis, the detailed process we have agreed regarding review of requests by the BMA and RCGP representatives, assessment of requests by DARS and independent scrutiny by IGARD, and the form, manner and period of the collection. The Data Provision Notice is the formal notice that is issued to general practices ahead of the data collection taking place.
- Business Rules. These provide a detailed technical specification of the data items and structure of this data collection and are available on our website

information about the Quality and Outcomes Framework, listed under Other Extracts – Emergency COVID-19 data collections.

- Direction. The COVID-19 Public Health Direction 2020, has been published on the NHS Digital website and is referenced in the Data Provision Notice and Transparency Notices. This ensures that NHS Digital is transparent in showing the public what it has been directed to do.
- Website content – a dedicated GDPPR webpage has been published to provide information on what the data will be used for, how it will be collected, who we will share data with and guidance for GPs as well as linking the user to other related sources of information. We will monitor the number of visits to the website content (hits, onwards clicks, downloads) and transparency content including this DPIA once published to gauge the level of engagement.
- We have consulted with groups that represent patients' and public interest in data including the Office of the National Data Guardian, Healthwatch England and the Information Commissioner's Office. We have consulted and shared our transparency materials and communications' content with them and asked for feedback on our communications and transparency approach, which we have incorporated. Additionally, we have sought feedback on our template General Practice Transparency Notice for GPs to use and sought feedback from two patient and public participation groups at Genomics England and HDRUK. As a result of this feedback, further work is underway on a revised Transparency Notice.
- Communications and Blogs – additional communications and blogs have been used to diversify information routes, such as the NHS Digital GP Bulletin, our Digital Transformation Blog on 22 May about why we are centralising GP data to support research during the pandemic and the RCGP COVID-19 Update Blog on 22 May which included GDPPR.
- Data Release Register – details of all disseminations of data made by DARS under its data sharing agreements will be published on its register of approved data releases.
- Minutes from IGARD Meetings – The minutes from each IGARD meeting are published within a week of the following IGARD meeting. We publish a rolling six months of meeting minutes.
- COVID-19 Response Information Governance Hub – examples of specific use of the GDPPR data will be published on the COVID-19 Response IG Hub to provide transparency over how the data is being used and is benefiting healthcare, treatment and public health.

5. How will you implement and support the rights of the individual in relation to this project?

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

- **The right to be informed** – Fair Processing information and Transparency Noticed for NHS Digital and GPs have been developed by NHS Digital as explained above and made available prior to the first data extract. .
- **The right of access** - An explanation about how an individual can request a copy of information that NHS Digital holds, including this GPES data, is published at: https://digital.nhs.uk/article/6851/How-to-make-a-subject-access-request. NHS Digital has established processes for handling Subject Access Requests through the Information Governance Team (IG Helpline Service). NHS Digital has established that a patient record can be extracted and the explanatory textual description supplied for each of the SNOMED codes. Any patent 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. 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 Collected Data will be supplied to NHS Digital every two weeks, any updates to the patient record will be supplied regularly.
- **The right to erasure** – an individual has the right to request erasure, however NHS Digital has the right to retain the data where necessary for legal purposes, such as defence of a legal claim and for audit purposes
- **The right to restrict processing** - Where an individual contests the accuracy of their personal data NHS Digital will consider the request.
- **The right to complain** - Individuals who believe that their data is not being processed in accordance with the law can complain to the Information Commissioner's Office (ICO)[9]. They can also contact NHS Digital's Data Protection Officer (DPO) regarding NHS Digital data processing activity and the DPO of their general practice regarding GP data processing activity. Details for the NHS Digital DPO are contained on the NHS Digital website in the General Transparency Notice.
- **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

[9] https://ico.org.uk/

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** – this right is applicable to processing based on the lawful basis of public task. However, NHS Digital would not share data following withdrawal of permission and would only process data following a withdrawal of permission for record keeping and legal purposes. This need would be considered compelling legitimate grounds which are likely to override the interests of the individual. This would be considered if the right was exercised.
- Patients that have registered a Type 1 objection with the GP Practice will not have their data shared with NHS Digital.
- Patients who have registered a National Data Opt Out will have their data shared with NHS Digital. NHS Digital will consider the application of the National Data Opt-Out on a case by case basis on any dissemination of identifiable patient data. This is because during this period of emergency, the National Data Opt-Out will not generally apply where data is used to support the coronavirus outbreak, due to the public interest in and legal requirements to share information. For example, the National Data opt Out does not apply where data is shared under the NHSD COPI Notice, which is a legal obligation. Nor does it apply for direct care. Notwithstanding this, DARS and IGARD will consider for all applications whether the National Data Opt Out should be applied where it would not otherwise prejudice the purpose of the processing. Although not related to this dataset, an example of where the NDOP was applied was in relation to sharing by NHS Digital of data with NHS Blood and Transplant for the Convalescent Plasma Trial, even although the NHS Digital COPI Notice applied to the dissemination.
- **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 Digital Information Governance Team and by DARS and IGARD on any application to access and use the data.

A summary of the application of individuals' rights is set out in the Table at Appendix D.

6. What measures do we have in place to ensure our processors comply with GDPR and our instructions in relation to this project?

No processors of NHS Digital are actively involved in the processing carried out on behalf of NHS Digital to collect, analyse or disseminate the data. The Collected Data is stored in DPS which uses Amazon Web Services (AWS), a cloud service hosted in the UK. AWS is a data processor for all data stored on DPS and NHS Digital has GDPR Article 28(3) compliant contract in place with AWS who have been appointed to provide the cloud services under Crown Commercial Services G-Cloud 9 contract. The contract is monitored by NHS Digital Commercial Team via a Key Account Manager. Any changes in our instructions to AWS would be processed in accordance with the change control

mechanisms under the contract.

PART E: RISK AND MITIGATION

5. Identification of the privacy and related risks[10] and mitigations[11]

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
1	**Ineffective or failed pseudonymisation at dissemination** If the data is disseminated in identifiable form when it was approved for access in pseudonymised form only, it would be a breach of the fairness principle 1 of GDPR. The risk to data subjects is that it could lead to unauthorised disclosure of their personal data including special category data.	Remote	Some	L	The data is pseudonymised through technical controls (including the removal of direct identifiers, minimisation of data by selection of specific codes, and creation of derived fields where appropriate). The data is also held under a data sharing framework contract and agreement, which restrict the use of the data to that which is set out within the agreement. The agreement also only allows other data to be linked with the GDPPR data where agreed within the DSA. The risk of reidentification is therefore minimised.	Reduced	L	Dave Roberts 09/09/20	

[10] Consider the potential impact on the individuals concerned and any harm or damage that might be caused by the processing involved in this project, in particular consider whether the processing could lead to an inability to exercise rights (including but not limited to privacy rights), an inability to access services or opportunities, loss of control over use of personal data etc. While there is no explicit definition of 'risk' in the GDPR, it's clear that it is about the risks to individuals interests, which may include risks to privacy and data protection rights, physical, material or non-material damage, discrimination, identify theft or fraud, financial loss, damage to reputation, loss of confidence, re-identification of pseudonymised data, security risks (including sources of risk and potential likelihood of breach and impact of each such breach) or more intangible harm like significant economic and social disadvantage, loss of public trust etc.

[11] A DPIA doesn't have to completely eradicate the risks altogether, but it should help to minimise the risks and assess whether or not the remaining risks are justified.

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	This risk to data subjects is likely to result in distress				Where data is held within NHS Digital's Trusted Research Environment (TRE) additional controls apply, in that record level data remains within the TRE by default. The exception is where specific permission has been given for the data to be extracted from the TRE to an individual organisation, at which point the same controls apply as for an extract.				
2	**Inadequate or no Transparency** Patients are not informed about how their personal data is being used through inadequate or lack of transparency material. The risk to data subjects is that their personal data is processed in an unfair or unexpected way, potentially causing	Remote	Some	**L**	See Section 2, Part D – 4 above. NHS Digital has made available a template GP Practice Privacy Notice and has published its own Transparency Notice Following patient consultation on the published Transparency Notices, NHS Digital is revising both the GP and the NHS Digital Transparency Notices to reflect feedback and improve the language and layering of the notices. Once implemented this	**Reduced**	**L**	Dave Roberts 09/09/20	Transparency Notices to be further updated – Susannah Strong leading with Richard Birmingham (IG) supporting. Revised Notices to be approved by DPO and Exec Director of IG before publication.

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	distress. This would breach the fairness principle 1 of GDPR and Articles 13 and 14. It could also result in data subjects not being aware of their data subject rights.				will further reduce this risk.				
3	**More data than is necessary for purposes specified in COVID-19 Direction are shared by GP Practices and Processed by NHS Digital.** Either because the data items were not all required or because the system for collection takes data items outside of the items covered by the DPN including Restricted Data. • The risk to data subjects is of distress	Remote	Some	**L**	Data will be extracted using GPES which will not pull all coded data but only the codes specified in the extract specification. Under the extraction arrangements carried out by the GP IT System Suppliers, the extractions are therefore limited to data fields that are detailed in the data specifications/ business rules. The specifications/business rules were developed in consultation with potential users of data to establish what data is necessary for their use cases and why. Legally Restricted Codes will be upheld (e.g. gender recognition)	**Reduced**	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	caused by processing data about them they did not expect would be shared with or by NHS Digital.				and excluded from the collection. Business rule has been written without the legally restricted codes and limited to SNOMED code sets. .				
4	**Patient opt-outs are not respected.** The consequences of not respecting the patient opt-out would cause data subjects distress **The risk to data subjects is that their expressed confirmation of the exercise of their opt-out right to have their data used for secondary**	Remote	Some	**L**	**Type 1 objections** will be upheld in collecting this data from General Practices and therefore the data for those patients who have registered a Type 1 objection with their GP will not be collected. The application of the **National Data Opt-Out** will be considered on a case by case basis for each dissemination and may or may not apply depending on the specific COVID-19 purposes for which the	Reduced	**L**	Dave Roberts 09/09/20	DARS to ensure that voluntary assessment of the application of NDOP is carried out and discussed with the data recipient at application stage. If NDOP is not applied, the explanation should be included in the application form. Owner: Garry Coleman

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	purposes (where such right applies) is not respected.				data is to be used. Application of the NDOP and exemptions will be assessed in accordance with published guidance – see Section 6.2 of the National Data Opt-Out Operational Policy Guidance. During this period of emergency, the National Data Opt-Out will not generally apply where data is used to support the coronavirus outbreak, due to the public interest in and legal requirements to share information. However, notwithstanding COPI Notices or other legal obligations on NHS Digital to share data, NHS Digital DARS will assess in conjunction with the data recipient at application stage whether the NDOP may be applied on a voluntary basis by NHS Digital. The assessment would consider the purposes for which the data is to be processed by the recipient and whether or not the NDOP would prejudice those purposes. For more information on the National Data Opt-Out and its				

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
					application during the COVID-19 period see.				
5	**Breach of security during the transmission to NHS Digital from GP System Suppliers, storage and processing of the Collected Data by NHS Digital resulting in the unauthorised disclosure of personal data.** **The risk to data subjects is that their personal data including special category data is disclosed to an unauthorised third party resulting in distress, anxiety and other harm to the rights and interests of the data subject**	Remote	Serious Harm	**M**	Data is transferred by GP System Suppliers to NHS Digital via MESH (https://digital.nhs.uk/services/message-exchange-for-social-care-and-health-mesh) MESH uses TLS mutual authentication (TLS v1 or above) to communicate between client side and Spine servers. There is client side mailbox password to control mailbox authentication. This is never passed on wire as HMAC (Hash-based message authentication) tokens are used on the basis of password, time stamp and random salt. HMAC token has limited time span and re-use is not possible as would be rejected by Spine (to stop man in the middle attacks). Data is encrypted in transit as all	**Reduced**	**L**	Dave Roberts 09/09/20	Functional and integration testing with each of the system suppliers who contribute to the end-to-end solution has been successfully completed and assured, confirming that data is accurately delivered to the MESH mailbox specified in data provision requests sent to the GPSS by the GPDC component of GPES. The GPES solution, which uses MESH as its underlying data transfer mechanism, has been fully solution assured by NHS Digital.

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
					interactions with the MESH service are through its REST API over HTTPS. Data is encrypted at rest with files being transferred over MESH being stored using AWS S3 server-side encryption using keys managed by NHS Digital within AWS KMS. Access to the S3 buckets is controlled via AWS IAM roles which are granted to services that process raw data for storage within the core DPS platform. Information including source and destination mailboxes, workflow identifiers are logged for all MESH transfers and the logs sent to the NHS Digital Splunk cloud service for storage outside the DPS infrastructure for audit and monitoring. The GPES solution, which uses MESH as its underlying data transfer mechanism, has been fully				

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
					solution assured by NHS Digital. GPES has a full-Service Management wrap. The Service Management team have access to the logs of all GPES messaging via dashboards and log querying tools, enabling visibility of the full data transfer pipeline. All GDPPR data landed onto DPS is schema-validated, which checks that the extract ID and structure of the returned data is valid. Additionally, analysts within the Primary Care Domain team perform analysis on the data each fortnight to check for discrepancies.				
6	**Breach of security by NHS Digital in the storage of the Collected Data in DPS resulting in the unauthorised disclosure of personal data to third parties and a personal data breach.**	Remote	Serius Harm	**M**	Data is encrypted at rest with files being transferred over MESH being stored using AWS S3 server-side encryption using keys managed by NHS Digital within AWS KMS. Access to the S3 buckets is controlled via AWS IAM roles which are granted to services that	**Reduced**	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	The risk to data subjects is distress, anxiety and other harm to the rights and interests				process raw data for storage within the core DPS platform. Information including source and destination mailboxes, workflow identifiers are logged for all MESH transfers and the logs sent to the NHS Digital Splunk cloud service for storage outside the DPS infrastructure for audit and monitoring.				
7	**Breach of security by NHS Digital in the dissemination of the Collected Data either through:** **(i) making the incorrect data available through the DAE; or** **(ii) the incorrect data available through a DARS data file** **resulting in the unauthorised disclosure of personal data to third**	Remote	Serious harm	**M**	Controlled process to convert an approved DARS request into access granted to defined Customer Views in DAE. ○ Data Sharing Agreement articulates specification to be provided to the customer ○ DMS data production have an assured processes for inputting DSA information to create a data view or extract conforming to the agreement. Testing and assurance of the Customer View development, to ensure appropriate data is included and pseudonymised where	**Reduced**	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	parties and a personal data breach. **The risk to data subjects is distress, anxiety and other harm to the rights and interests**				necessary ○ The extract method has been tested by the assurance team proving that given the correct inputs the correct customer outputs are produced ○ There is a sign-off process for individual customer views/extracts whereby a senior manager compared the data produced with the specification in the DSA to ensure these are aligned Controlled process for removing data should any data breach be identified ○ Technical process exists to remove customer access should a breach be identified ○ Process in place for reporting any data breaches through the appropriate channels via National Service Desk				

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
8	**Risk of unauthorised access of the Collected Data within NHS Digital resulting in the data being used for unauthorised purposes** **The risk to data subjects is that their personal data including special category data is processed by unauthorised third party resulting in distress, anxiety and other harm to the rights and interests of the data subject** This risk to data subjects could result in: • Patient complaints • ICO regulatory action • Patient compensation claims	Remote	Serious harm	**M**	All analysis carried out internally within NHS Digital under the COVID-19 Direction needs to be approved by NHS Digital's Information Governance Team and Caldicott Guardian, with advice sought from IGARD where appropriate. This process involves the review of the analysis proposal, the data required, how minimisation principle is complied with and ensures there is a legal basis for the analysis. Any substantial analysis will require a DPIA to be carried out. Access by analysts to the data is governed by the Clear Data Access process as explained above in 'Section 2 Part C – 3, 4, Internal Parties', and will only be granted to analysts who are carrying out approved NHS Digital analysis following IG and Cg approval above, or carrying out approved analysis for a DARS customer pursuant to an approved DARS DSA following request by	**Reduced**	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
					DARS. Access to the Collected Data by internal analysts is controlled by the completion of COVID-19 access forms, which must be approved by NHS Digital Information Governance function. This access is strictly limited and subject to authorization by the Information Asset Owner through NHS Digital's own Clear Data Access internal approval process.				
9	**Risk of data being disseminated to organisations that do not meet the required security standard expected by NHS Digital** **The risk to data subjects is that their personal data including special category data is subject to a security breach and disclosed to an unauthorised third party**	Remote	M	**L**	Part of the DARS application process includes checking that the organisations that control/process the data have appropriate safeguards in place for secure handling of the data and they meet the obligations in their data sharing contract and Data Sharing Agreement. By default, organisations are expected to have satisfactorily completed the Data Security and Protection Toolkit (DSPT) with respect to the teams accessing	**Reduced**	**L**	Dave Roberts 09/90/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	resulting in distress, anxiety and other harm to the rights and interests of the data subject				the data. Where organisations do not have DSPT but have alternative security controls such as ISO or a System Level Security Policy, these may be accepted if deemed acceptable by NHS Digital's IT Security team. Further detail is available through the Data Access Request Service area of the NHS Digital website, which the appropriate standard for access is listed,				
10	**The Collected Data are used for a purpose outside of what is defined in the COVID-19 Direction by data recipients.** The risk to data subjects is that their data could be processed for purposes beyond those which they expected and which were set out in the Transparency Notices	Remote	M	**L**	The DARS application process assesses how the Collected Data is used and an application will only be accepted for COVID-19 planning and research purposes. The DARS Data Sharing Contract for the dissemination of the Collected Data will set out the purposes for which the data can be used and that the data can't be used for any other purposes, without applying to DARS for a change to the Agreement. DARS publishes details of those with whom it has shared data and the	**Reduced**	**L**	Dave Roberts 09/09/20	DARS Audit Plan to include audit of high risk GPES data recipients. Owner: Garry Coleman

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	provided by NHS Digital and GPs This could cause the data subject to suffer distress and otherwise impact on their rights and interests including that their data is used for purposes that they might reasonably consider inappropriate or unethical.				purposes for this in the Data Release Register. NHS Digital have rights to audit agreement compliance and take action if there are breaches of the agreement. NHS Digital will proactively audit higher risk data use cases/recipients of data as part of its rolling DARS Data Audit Plan.				
11	**The personal data is processed for longer than necessary for the purposes set out in the COVID-19 Direction by** **(i) NHS DIgital** **(ii) the data recipient** Retaining and processing data for longer than is necessary will cause distress to data recipients.	Remote	M	**L**	The Collected Data will be kept for 8 years from the date it is last collected for legal record keeping reasons as explained above. This retention period is in line with the NHS Digital Records Management Policy and the NHS Records Management Code. Responsibility for ensuring retention periods are complied with rests with the Information Asset Owner (IAO). IAOs must review all	**Reduced**	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	.				of their Information Assets annually and confirm the Unified Register is up to date as part of the annual NHS Digital Data Security and Protection Toolkit submission. NHS Digital has created a separate area within DPS where COVID-19 data sets are all stored and/or processed under the COVID-19 Direction for ease of identification. When data is to be destroyed, it will be destroyed securely with the decision and action recorded in accordance with our Records and Document Management Policy and Corporate Retention and Disposal Framework: Implementation Process. For data stored on DPS the process for destruction is to carry out a delete action on the relevant data, this takes 180 days for complete destruction as per AWS terms and conditions.				

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
					(ii) Data Recipients Data recipients must specify their intended data processing and retention period in their DARS application, which will relate to the purposes for which they are to process the data. The retention period is specified in the Data Sharing Agreement and a DSA will remain in place for the duration the data is to be retained. The DSA will require secure destruction at the end of the permitted period and provision of a signed data destruction certificate from the DPO of the data recipient as evidence the data has been securely destroyed. The appropriateness of the retention period is assessed as part of the DARS process				

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
12	**Risk that inaccurate data is collected, analysed and disseminated to authorised organisations** Data will be collected as an initial bulk extract and then on a fortnightly basis. Due to the 2-week extraction frequency, there is a risk that the data may not be up to date at the time of dissemination. . This could lead to inaccurate data being used for authorised purposes. As the data is being used for secondary use purposes it is unlikely that this will have an impact on the individual concerned.	Remote	Minimal	**L**	Responsibility for data accuracy within patient records lies with the GP Practices as the source data controller. Where updates are made to GP records, these updates will be collected by NHS Digital as part of the next data collection. The data collected by NHS Digital is therefore updated every 2 weeks. NHS Digital shall ensure the last date of the collection is clearly known by the recipients of the data (e.g. time stamp upon data extraction) to prevent misinterpretation of the data recency. Data extraction frequency will be reviewed and increased when there is clear requirement and capacity is available to execute increased frequency. For the overall quality of the data, when the Collected Data is received by NHS Digital, it is	Reduced	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
					schema-validated, which checks that the extract ID and structure of the returned data is valid. Additionally, analysts within the Primary Care Domain team perform analysis on the data each fortnight to check for discrepancies.				
13	**Risk that Collected Data is transferred outside of the UK to a jurisdiction which does not provide for sufficient protections of the rights of data subjects as required under GDPR** **The risk to data subjects is that the data protection arrangements in the jurisdiction where processed are not as robust as in the UK and do not provide the same level of safeguards over their data or afford them the ability to exercise**	Remote	Minimal	**L**	The geographical location of the data once transferred to NHS Digital data remains within the UK jurisdiction in the DPS Platform which is hosted in the AWS cloud within the UK. Data recipients must specify the location of processing and location of the storage of data in their DARS application. DARS will assess the adequacy of the location as part of the application process, which will be overseen by IGARD. Adequacy will include an assessment of how any overseas processing or storage complies with GDPR. Location of processing and storage of data is included in	**Reduced**	**L**	Dave Roberts 09/09/20	

No	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 L/M/H	Mitigations of the risk	Effect on risk (Eliminated/ Reduced/ Accepted)	Residual risk L/M/H	Measure approved (Name and Date)	Actions to be taken & taken (Date and responsibility for completion)
	their data subject rights in the same way.				the Data Sharing Agreement and can be subject to audit				

PART F: APPROVAL OF DPIA AND RISKS

Please complete the table below to show who has approved the privacy risks and the DPIA on behalf of NHS Digital.

Name of Approver	Position	Date	Signature
Dave Roberts	Information Asset Owner	10/11/2020	

APPENDIX A - GLOSSARY

Terms

Term	Definition
Automated Decision Making	A decision which has a significant effect on an individual and is taken solely on the basis of automated Processing of Personal Data. This means processing using, for example, software code or an algorithm, which does not require human intervention.
Bucket	In cloud computing buckets are the basic containers that hold data. Everything that is stored in Cloud Storage must be contained in a bucket and they can be used to organize data and control access to data.
Data Access Request Service	The Data Access Request Service (DARS) can offer clinicians, researchers and commissioners the data required to help improve NHS services. It is a function of NHS Digital.
Data View	Data views are security-controlled data assets within the DAE as our default method of dissemination for GDPPR data. Data views are pre-existing queries, which provide a user with access to a constrained version of the underlying data. Each data view is a restricted window on the underlying data, not a copy of it. NHS Digital creates specific views of data within the DAE for customers, which reflect what the customer is approved to access in their Data Sharing Agreement, depending on the DARS-authorised access requirements for a particular applicant.
Denial of Service	Where decisions are made about an individual's access to a product, service, opportunity or benefit which will be based to any extent on **Automated Decision Making** (including **Profiling**) or involves the processing of **Special Category Data.**

Term	Definition
Data Extract	Data that is collected from the GP IT systems by GPES over a specified period of time, limited to data fields that are detailed in the data specifications/ business rules. Appendix B illustrates the component parts of GPES involved to produce and flow the GPES Data for Pandemic Planning and Research extract.
Invisible Processing	Processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with GDPR Article 14 would prove impossible or involve disproportionate effort.
Personal Data	Any information relating to an identified or identifiable individual (an "individual" or "data subject").
Processing	Any operation performed on Personal Data.
Profiling	Automated processing of Personal Data in order to evaluate certain personal aspects relating to an individual, in particular to analyse or predict aspects concerning that person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.
Pseudonymisation	Means the Processing of Personal Data in such a manner that the Personal Data can no longer be attributed to a specific data subject without the use of additional information, which is held separately.
Release of Data	Dissemination of data upon the approval of data access request by DARS/IGARD and signing of data sharing agreements between the data requestor and NHS Digital.

Term	Definition
Special Categories of Personal Data	Personal Data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purposes of uniquely identifying an individual, data concerning health, or data concerning a natural person's sex life or sexual orientation.
Systematic Monitoring	Processing of personal data in order to observe, monitor or control data subjects in a systematic (i.e. ordered, predetermined or regular) way.
Vulnerable Individuals	For the purposes of health data this is taken to mean Children and others subject to safeguarding activity.

Abbreviations

Abbreviation	Meaning	Explanation/Link
AWS	Amazon Web Services	https://aws.amazon.com/what-is-aws/
BMA	British Medical Association	https://www.bma.org.uk/
CCG	Clinical Commissioning Group	https://www.nhscc.org/ccgs/ https://www.england.nhs.uk/ccgs/
CDA	Clear Data Access	The CDA process is an internal process to control access to data by NHS Digital staff. Staff must apply for access via a form which must be approved by line managers and the relevant Information Asset Owner to grant access. Staff will only be granted access to those functions or data required to perform their role and access to identifiable data is only granted where there is a specific need that cannot be met using de-identified data. This will be assessed as part of the approval for access to data for internal NHS Digital analysis or as part of the approval for dissemination of a data to a data recipient through DARS.
DAE	Data Access Environment	The Data Access Environment (DAE) presents the data stored within the DPS platform. It is the secure way users can remotely access NHS Digital data, including linked information, while ensuring the right person, with the right permissions gets the right data, in accordance with their Data Sharing Agreement (DSA). https://digital.nhs.uk/services/data-access-environment-dae
DARS	Data Access Request Service	https://digital.nhs.uk/services/data-access-request-service-dars
DHSC	Department of Health and Social Care	https://www.gov.uk/government/organisations/department-of-health-and-social-care/about

Abbreviation	Meaning	Explanation/Link
DPIA	Data Protection Impact Assessment	https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/
DPN	Data Provision Notice	https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns
DPS	Data Processing Services	Data Processing Services is a term used to both describe generally the secure technologies and processes used by NHS Digital to enable us to collect, process and access data, and specifically within this DPIA to the platform and series of processes for receiving and transforming national data collected from care provider and other care related organisations. https://digital.nhs.uk/data-and-information/data-insights-and-statistics/improving-our-data-processing-services
DSA	Data Sharing Agreement	https://ico.org.uk/media/for-organisations/documents/1068/data_sharing_code_of_practice.pdf
COPI Notice	Notices issued under Regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002. This includes the Notices issued to NHS Digital by the Secretary of State for Health and Social Care dated 17 March 2020 and 29 July 2020	https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information
GDPR	General Data Protection Regulation	https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/

Abbreviation	Meaning	Explanation/Link
GDPPR	GPES Data for Pandemic Planning and Research	https://digital.nhs.uk/coronavirus/gpes-data-for-pandemic-planning-and-research
GPITF	GP IT Futures	https://digital.nhs.uk/services/future-gp-it-systems-and-services
GPES	GP Extraction Service	The General Practice Extraction Service (GPES) collects information for a wide range of purposes, including providing GP payments. It works with the Calculating Quality Reporting Service (CQRS) and GP clinical systems as part of the GP Collections service. GPES collects data automatically over a specified period of time - this is known as an extract. Appendix B illustrates the component parts of GPES involved to produce and flow the GPES Data for Pandemic Planning and Research extract. https://digital.nhs.uk/services/general-practice-extraction-service
GPSS	GP IT System Supplier	GPES collects data from four principal GP IT system suppliers including TPP SystmOne, EMIS Web, InPS Vision and Microtest Evolution.
IAO	Information Asset Owner	https://www.gov.uk/government/publications/information-asset-owner-role-guidance
IGARD	Independent Group Advising on the Release of Data	https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/independent-group-advising-on-the-release-of-data
JGPITC	Joint GP IT Committee	Together the RCGP and BMA form the Joint GP IT Committee, which is a contractually-mandated committee representing the views of GPs from all 4 nations and users of all GP systems in discussions relating to the use and management of these systems and GP data created therein
LSOA	Lower Layer Super Output Area	https://www.datadictionary.nhs.uk/data_dictionary/nhs_business_definitions/l/lower_layer_super_output_area_de.asp?shownav=1

Abbreviation	Meaning	Explanation/Link
MESH	Message Exchange for Social Care and Health	MESH is a service provided by NHS Digital and is the main secure large file transfer service used across health and social care organisations for clinical and other data. https://digital.nhs.uk/services/message-exchange-for-social-care-and-health-mesh
PHE	Public Health England	https://www.gov.uk/government/organisations/public-health-england/about
RCGP	Royal College of General Practitioners	https://www.rcgp.org.uk/about-us.aspx
SEFT	Secure Electronic File Transfer	NHS Digital business teams can use the Secure Electronic File Transfer (SEFT) to transfer data securely to and from any external organisation. https://digital.nhs.uk/services/transfer-data-securely
SLSP	System Level Security Policy	The SLSP is a trustworthy, concise and considered view of the Information Security of a system at any given time. This information is used to fulfil NHS Digital's obligation to current legislation on appropriate technical controls to safeguard its information
SoS	Secretary of State for Health & Social Care	https://www.gov.uk/government/ministers/secretary-of-state-for-health-and-social-care
SPOC	NHSX Single Point of Contact COVID-19 request service	Alternative name for the NHSX Single Triage Service to request access to health and care data in order to support the COVID-19 response. https://www.england.nhs.uk/ourwork/tsd/data-info/

APPENDIX B – Data Flow Diagram

GPES Data for Pandemic Planning and Research Data Flow

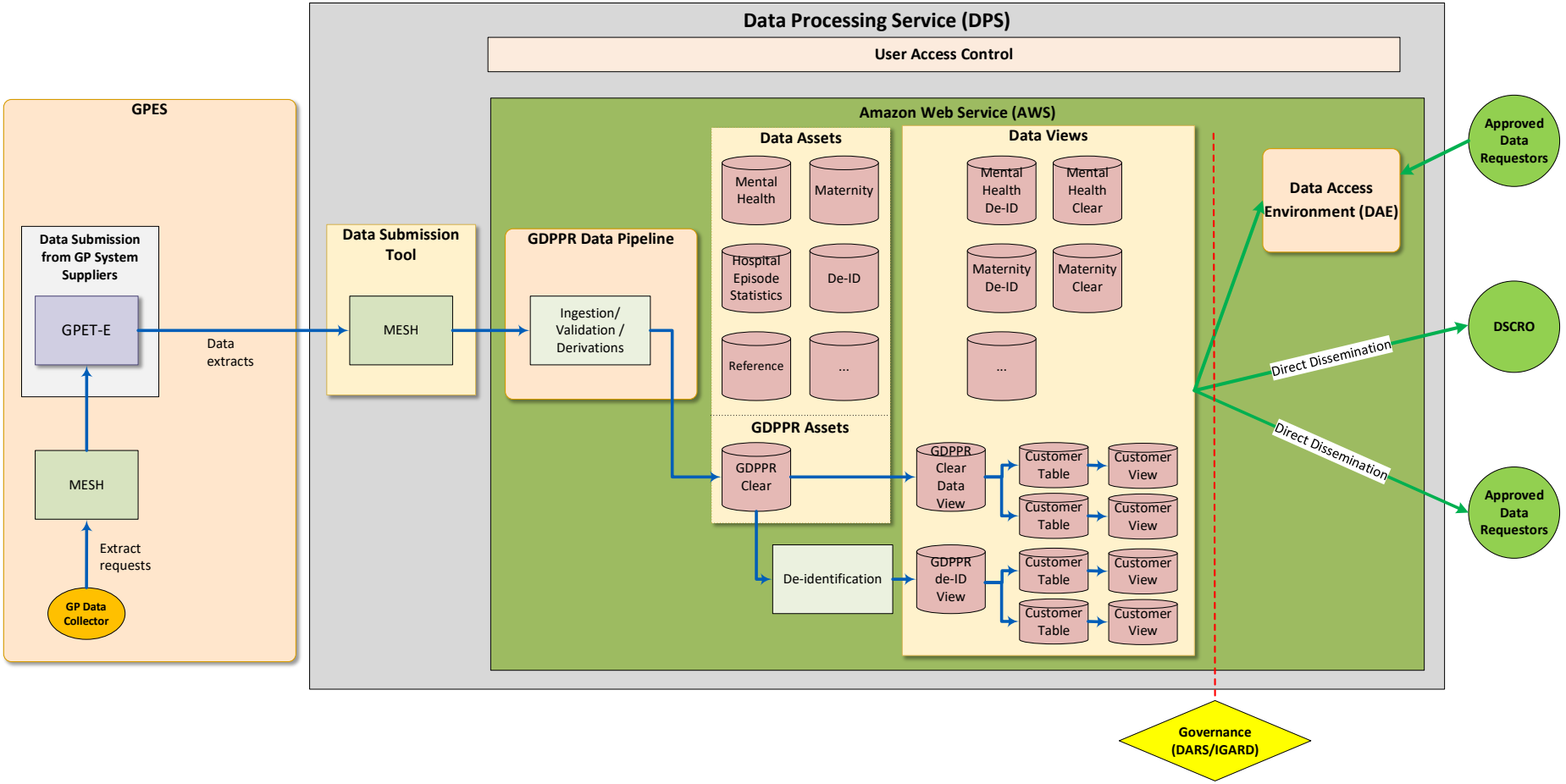

Figure 2: Overview of Data Processing Service

APPENDIX C – Data Processing Service and Data Access Environment

The Data Processing Service (DPS) and Data Access Environment (DAE) are fundamental components of the data processing, storage and dissemination for GDPPR and there are frequent references to both throughout this DPIA. Information is provided in this section for ease of understanding and referencing.

DPS & DAE Overview

What is the DPS?

Within this DPIA Data Processing Service (DPS) refers to the single platform and series of processes for receiving and transforming national data collected from care provider and other care related organisations, and presenting it for internal and external use.

What is the DAE?

The DAE is a term used to describe an area within the DPS platform which presents the data stored within DPS. It is the secure way users can remotely access better linked information, while ensuring the right person, with the right permissions gets the right data, in accordance with their Data Sharing Agreement (DSA).

DAE provides a single access environment for internal and external users to access this data and supports a number of presentation tools. By default, users cannot download the results of queries from DAE. However, there are cases where this is necessary in which case the user is granted specific permission to download data.

DPS Security

The DPS infrastructure is hosted on AWS Public Cloud and has been designed around the good practice guidance and risk model described in NHS and social care data: off-shoring and the use of public cloud services[12] written jointly by NHS Digital, NHS England, the Department of Health and Social Care and NHS Improvement.
Due to the nature of data held within DPS, the service is within the highest category of risk within the risk model.

Organisation controls for internal access to DPS data

All NHS Digital staff undertake mandated Information Governance training that includes how the data should be handled and shared appropriately, what to do when a personal

[12] https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/nhs-and-social-care-data-off-shoring-and-the-use-of-public-cloud-services

data breach has been identified, and awareness of policies e.g. Data Protection Legislation policy, Bring Your Own Device (BYOD) policy.

All staff requiring access to identifiable data must have national security vetting Security Check level clearance.

All NHS Digital staff to the DPS and DAE systems is controlled through Clear Data Access (CDA) forms approved by line managers and the IAO for the dataset. Staff will only be granted access to functions or data required to perform their role. Access to identifiable data is only granted where there is a specific need that cannot be met using de-identified data.

Organisation controls for external access to DPS data

Organisations and individuals wanting access to data held in DPS must apply for access via the DARS process[13]. For GDPPR data requests much be made via the NHSX Single Triage Service for triage and subsequent handover to DARS of appropriate requests. If approved, the user will be given access to the requested data via DAE (excepting direct dissemination via delivery of a data file, see above). Access is restricted to only the approved data.

User Authentication

DPS incorporates a Single-Sign-On service for the management of service user identities. This is built on the supported and widely used Open-source Identity Provider software KeyCloak. This is configured to enforce Multi Factor Authentication.

Data in Transit

Data in transit is always sent over encrypted connections either via:

- HTTPS
- SFTP / SCP

All authentication is carried out over encrypted links

Unencrypted traffic is sometimes used for some non-Personal data such as dashboards. This is being phased out.

Data at Rest

Data sets are stored as objects in AWS S3 Buckets with access controlled through the AWS Identity and Access Management (IAM) mechanisms.

The buckets are not publicly readable. This is verified by regular security monitoring and

[13] See https://digital.nhs.uk/services/data-access-request-service-dars

auditing by the NHS Digital AWS service.

The data is encrypted at rest in S3 using AES-256 and keys maintained by NHS Digital and stored in AWS KMS AWS Key Management Service (KMS). For data to be read, the service or user must have read access on the object **and** on the associated key.

S3 Pre-signed URLs are used to provide temporary upload and download access

Different S3 buckets are used for data at different levels of sensitivity and at different points in the processing pipeline.

Data Access Control

The data access services that present data to service users are associated with AWS service principal(s). The principal(s) are granted permissions to some subset of the underlying S3 objects.

On top of this, user level access control based in identities from the DPS Identity Provider (KeyCloak) and permissions are enforced by the Databricks Operation Security features.

Data access is usually granted at the level of a database schema rather than at table or column level. Schemas intended for service users will consist exclusively of database views designed to included only the data that that individual user is entitled to see.

Data Privacy

Personal Identifiable Data such as NHS Numbers and identifiers returned by the NHS Digital Master Person Service is processed by the NHS Digital run instance of the Privitar Data Privacy Service to produce a series of Pseudonyms representing the data item across a set of privacy domains.

Where a user is not entitled to see Personal Identifiable data, but is entitled to view pseudonymised data, the views used for access control will map the identifiers to the appropriate domain pseudonym.

Data Analysis Tooling and Data Output

DPS provides analytical tooling including the HUE SQL Workbench and the Databricks Notebook User Interface. There are plans to extend this tooling to include other popular workbench tools such as R-Studio.

To reduce the risk of sensitive data being taken out of the secure environment, access to the Production Data Processing Service analytical tool is provided through Apache Guacamole - which offers a remote desktop running in the user's web browser. Restrictions on copy-and-paste are in place.

Users can be granted permission to send some categories of data - including CSV files - to a data output service which will deliver to an S3 bucket protected by AWS Identity Management linked to the KeyCloak supplied identity.

Systems Administration Access Control

Systems Administration is typically based on the Bastion model - described in The National Cyber Security Centre Systems Administration Architectures[14]. The risks associated with this approach are mitigated through the user of network access control to limit connections to hosts on the Health and Social Care Network (HSCN).

[14] https://www.ncsc.gov.uk/guidance/systems-administration-architectures

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

- **A summary of the rights that apply in relation to the legal bases used is provided below. They are separated into a list of applicable rights in respect of the collection and analysis of the personal data and dissemination of the personal data:**

Collection and analysis

Rights	Right available where we are relying solely on public task (Article 6(1)(e) of GDPR above)?	Right available where we are relying on Legal Obligation – by virtue of COVID-19 Public Health Direction under S254 of 2012 Act (Article 6(1)(c) of GDPR above)?	Right available where processing is pursuant to the exemption contained within Article 9(2)(g) of the GDPR – substantial public interest?
Be informed	✓	✓	✓
Get access	✓	✓	✓
Rectify or change	✓	✓	✓
Erase or move	✓	✓	✓
Restrict or stop processing	✓	✗	✓

Rights	Right available where we are relying solely on public task (Article 6(1)(e) of GDPR above)?	Right available where we are relying on Legal Obligation – by virtue of COVID-19 Public Health Direction under S254 of 2012 Act (Article 6(1)(c) of GDPR above)?	Right available where processing is pursuant to the exemption contained within Article 9(2)(g) of the GDPR – substantial public interest?
Move, copy or transfer	✗	✗	✗
Object to processing or use	✓	✗	✓
Know if a decision was made by a computer rather than a person	✓	✓	✓
Raise a concern	✓	✓	✓

Dissemination

Rights	Right available where we are relying solely on public task (Article 6(1)(e) of GDPR above)?	Right available where we are relying on Legal Obligation – by virtue of COVID-19 Public Health Direction under S254 of 2012 Act (Article 6(1)(c) of GDPR above)	Article 6(1)(d) Right available where processing is necessary in order to protect the vital interests of the data subject or others?	Article 6(1)(f) Right available where processing is necessary for the purpose of legitimate interests?	Article 9(2)(I) Right available where processing is necessary for achieving the public interest in the area of public health?	Article 9(2)(h) Right available where processing is necessary for the purpose of medical diagnosis, the provision of health and social care treatment or the management of health and social care systems?	Article 9(2)(j) Right available where we are processing for scientific research and statistical purposes in relation to anonymous statistics?	Article 9(2)(g) Right available where we are processing on the basis of substantial public interest?
Be informed	✓	✓	✓	✓	✓	✓	✓	✓
Get access	✓	✓	✓	✓	✓	✓	✗	✓
Rectify or change	✓	✓	✓	✓	✓	✓	✗	✓
Erase or move	✓	✓	✓	✓	✓	✓	✓	✓
Restrict or stop processing	✓	✗	✗	✓	✗	✗	✗	✗

Rights	Right available where we are relying solely on public task (Article 6(1)(e) of GDPR above)?	Right available where we are relying on Legal Obligation – by virtue of COVID-19 Public Health Direction under S254 of 2012 Act (Article 6(1)(c) of GDPR above)	Article 6(1)(d) Right available where processing is necessary in order to protect the vital interests of the data subject or others?	Article 6(1)(f) Right available where processing is necessary for the purpose of legitimate interests?	Article 9(2)(I) Right available where processing is necessary for achieving the public interest in the area of public health?	Article 9(2)(h) Right available where processing is necessary for the purpose of medical diagnosis, the provision of health and social care treatment or the management of health and social care systems?	Article 9(2)(j) Right available where we are processing for scientific research and statistical purposes in relation to anonymous statistics?	Article 9(2)(g) Right available where we are processing on the basis of substantial public interest?
Move, copy or transfer	✗	✗	✓	✓	✗	✓	✓	✓
Object to processing or use	✓	✗	✗	✓	✗	✗	✗	✗
Know if a decision was made by a computer rather than a person	✓	✓	✓	✓	✓	✓	✓	✓

Rights	Right available where we are relying solely on public task (Article 6(1)(e) of GDPR above)?	Right available where we are relying on Legal Obligation – by virtue of COVID-19 Public Health Direction under S254 of 2012 Act (Article 6(1)(c) of GDPR above)	Article 6(1)(d) Right available where processing is necessary in order to protect the vital interests of the data subject or others?	Article 6(1)(f) Right available where processing is necessary for the purpose of legitimate interests?	Article 9(2)(I) Right available where processing is necessary for achieving the public interest in the area of public health?	Article 9(2)(h) Right available where processing is necessary for the purpose of medical diagnosis, the provision of health and social care treatment or the management of health and social care systems?	Article 9(2)(j) Right available where we are processing for scientific research and statistical purposes in relation to anonymous statistics?	Article 9(2)(g) Right available where we are processing on the basis of substantial public interest?
Raise a concern	✓	✓	✓	✓	✓	✓	✓	✓

- The **right to be informed** – Fair Processing information (Privacy Notice) and a Transparency Notice have been developed by NHS Digital as explained above and published prior to the collection of GPES Data for Pandemic Planning and Research.
- The **right of access** - Individuals can request access to information by making a Subject Access Request to NHS Digital. An explanation about how an individual can request a copy of information that NHS Digital holds is published at: https://digital.nhs.uk/article/6851/How-to-make-a-subject-access-request. NHS Digital has established processes for handling Subject

Access Requests. Information about the right to make a Subject Access Request and how to do it are also contained within the published Privacy Notice referred to above.

- The **right to rectification** - The right for individuals to have inaccurate personal data rectified, the data are derived from GP systems and are updated every two weeks. This means that unless the corresponding GP record has been rectified the same data will be collected again.
- the **right to erasure** –an individual has the right to request erasure, however NHS Digital has the right to retain the data where necessary for legal purposes, such as defence of a legal claim and for audit purposes
- the **right to restrict processing** – this will apply
- the **right to data portability**- is not applicable some processing because under article 20 (3) the processing is being carried out on the basis of public task or the exercise of official authority vested in the controller
- the **right to object** – this right is applicable to processing based on the lawful basis of public task. However, NHS Digital would not share data following withdrawal of permission and would only process data following a withdrawal of permission for record keeping and legal purposes. This need would be considered compelling legitimate grounds which are likely to override the interests of the individual. This would be considered if the right was exercised.
- the **right to raise a concern** with NHS Digital and the Information Commissioner's Office at any time

Individuals can exercise their rights, or find out more information about this service, and how it processes personal data by contacting NHS Digital and

making a request. This includes any request to obtain a copy of their personal data. Any request to exercise a right listed above or to obtain a copy of their personal data can be made at the web page; https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/publication-scheme/how-to-make-a-subject-access-request

Individuals who believe that their data is not being processed in accordance with the law can complain to the Information Commissioner's Office (ICO). They can also contact NHS Digital's Data Protection Officer (DPO) regarding any NHS Digital data processing activity. Details for the NHS Digital DPO are included in the Privacy Notice

Individuals will be made aware of their rights through the following documents hosted on the service that will be available from the footer of every page in the service:

- Privacy Notice (Transparency Notice) - already published

