

Greater Manchester Information Governance Group (GMIGG)¹

Data Protection Impact Assessment (DPIA)

The instrument for a privacy impact assessment (PIA) or data protection impact assessment (DPIA) was introduced with the General Data Protection Regulation (Art. 35 of the GDPR). This refers to the obligation of the controller to conduct an impact assessment and to document it before starting the intended data processing. Article 35(1) of the General Data Protection Regulations says that you must do a DPIA where a type of processing is likely to result in a high risk to the rights and freedoms of individuals:

"Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks."

The DPIA Process

The Data Protection Act is mainly concerned with the disclosure of personal data outside the data controller's own boundaries.²

If the data is to be **anonymised PRIOR** to any processing you **may** not need to complete this DPIA and should review:

- question 1.20
- section 2

and liaise with your IG Lead to confirm completion is not required.

Otherwise:

- 1) Please complete each section 1 - 4 with as much detail as possible. Your IG lead can complete section 5 but may need additional information from you. Section 6 onwards can be completed together with your IG Lead.
- 2) Once you submit the DPIA for approval to/via your Information Governance Lead/Data Protection Officer (DPO)
 - a. The DPIA proforma will be vetted and you may receive some comments / questions asking for further information. Please answer these promptly and resend the DPIA again.
 - b. The DPIA then goes for approval. It is considered for approval by the relevant IG internal approval process.
- 3) Once approved, the process / system can start to be introduced or modification to an existing system / process can continue.
- 4) **If you proceed with the initiative without completing the DPIA and without approval via the IG DPIA approval process, you are putting the organisation at risk of being in breach of the DP legislation which may result in disciplinary procedures being invoked.**

Initiative/System/ Process name:	NHS GM Secure Data Environment – Analytics and Data Science Platform (ADSP)
Link to any wider initiative: (if applicable)	<ul style="list-style-type: none">• Secure data environment for NHS health and social care data - policy guidelines Published 6 September 2022 and updated 22 Dec 2022• Data saves lives: reshaping health and social care with data – updated June 2022• Better, broader, safer: using health data for research and analysis – published 7 April 2022• NHS England 2022/23 priorities and operational planning guidance – Version 3 – updated February 2022• NHS Integrated Care Systems Design Framework – June 2021

¹ GMIGG is one of the regional Strategic Information Governance Networks (SIGN) groups that feed into the national SIGN supported by NHS England and NHS Digital.

² [ICO – Anonymisation code](#)

	<ul style="list-style-type: none"> • NHS What Good Looks Like framework – published August 2021 and updated October 2021
Date Initiative due to go live/commenced:	The ADSP has been under development for several years within GM. This DPIA takes into account the emerging national Secure Data Environment policy and the National Data Guardian requirements regarding the processing of data from Shared Care Records for secondary uses and research.
Date DPIA commenced:	07/11/2022

DPIA Contact Details: *Please list all main contacts involved in completing the DPIA including relevant service lead*

Name	Role	Organisation/ dept.	Email
Jenny Spiers	Head of IG	Health Innovation Manchester	Jenny.Spiers@healthinnovationmanchester.com
Graham Hayler	Deputy Director of Data and Analytics (Place and Data Governance)	NHS GM Integrated Care	graham.hayler@nhs.net
Graham Beales	Deputy Director of Data and Analytics (Hospital Care, Corporate performance, and Infrastructure)	NHS GM Integrated Care	graham.beales@nhs.net
Matt Hennessey	Chief Intelligence and Analytics Officer	NHS GM Integrated Care	matt.hennessey@nhs.net

Version	Date	Amendment History
0.1	7 November 2022	1 st draft
0.2	2 May 2023	Revised draft
0.3	27 June 2023	Updated draft
0.4	7 August 2023	Updated draft
0.5	11 August 2023	Updated draft
0.6	17 August 2023	Expiry date of Tableau contract extended
1.0	31 st October 2023	Version updated to reflect SIRO approval
1.1	15 th April 2024	Governance Groups updated, and inclusion of the Patient Re-Identification (appendix F) to support direct care. Section 251 CAG support included and ICO registration renewal updates for data processors updated
1.2	2 nd January 2025	Updated to reflect the move Tableau Cloud and amendments to the User Management processes and software (see Appendix 3)

DPIA reviewers	Date(s)	Version(s)
Emilie Yates – V-lex	March, May, June/July	0.1, 0.2, 0.3
NHS GM DPOs – CC/SP	17 August 2023	0.5, 0.6
NHS GM SIRO	31 st October 2023	0.6
Deputy Director of Data and Analytics (Place and Data Governance)	27 th September 2024	1.1

DPIA approval	Date	Version(s)
NHS GM IGOG	24 August 2023	0.6 – final version for discussion. Endorsed by NHS GM Data Protection Officers.
NHS GM SIRO	31 st October 2023	1.0 (email approval from SIRO)

Definitions	
-------------	--

Analytics and Data Science Platform - ADSP	The GM Insight platform was created as part of the GM Local Health and Care programme. It consists of best of breed storage and processing tools hosted by Arden and Gem CSU within their Microsoft Azure tenancy. All cloud hosted assets meet the stringent security and access control requirements for sensitive health data and have been assured by NHS Digital (now NHSE) with annual security reviews conducted by an independent IT security provider.
Arden & GEM (AGEM CSU)	ARDEN AND GREATER EAST MIDLANDS COMMISSIONING SUPPORT UNIT - an arm's length service organisation of NHS Commissioning Board engaged by NHS GM to provide services in relation to clinical commissioning support including DSCRO.
Anonymised data	Data in a form that does not identify individuals and where identification through its combination with other data is not likely to take place.
CAG	Confidentiality Advisory Group - an independent body which provides expert advice on the use of confidential patient information. This includes providing advice to the Health Research Authority (HRA) for research uses. It also provides advice to the Secretary of State for Health for non-research uses. https://www.hra.nhs.uk/about-us/committees-and-services/confidentiality-advisory-group/
DSCRO (NHS Digital)	A Data Services for Commissioners Regional Office (DSCRO) work with data from a wide range of sources including NHS hospital trusts, Local Authorities and GP Practices in regional processing centres. Staff follow strict rules on accessing, analysing and processing data. The powers are granted to the organisation by the Health and Social Care Act 2012 which means that staff are operating within an approved legal framework. Whilst staff within the DSCROs are employed by the CSUs they are seconded to NHS England, and must have appropriate approvals in place to allow them to access, handle and process identifiable data. They adhere to the same strict data security policies and controls as permanent staff. See link here for further information. Greater Manchester utilises the service via the AGEM CSU.
GM	Greater Manchester
GMCA	Greater Manchester Combined Authority
GMCR	Greater Manchester Care Record – see link here
National Data Opt Out (NDOO)	National Data Opt Outs apply to all organisations providing or coordinating publicly funded health or adult social care in England. It prevents the sharing of identifiable patient data from those organisations and NHS England for reasons other than individual direct care (subject to exemptions such as legal obligations and public health). This opt out can be completed via the NHS Digital website. More information about the National Data Opt-out and how the data is used can be found here: https://www.nhs.uk/your-nhs-data-matters/ . The NDOO does not apply to data processed under the terms of the COPI Notice or where the data is anonymised in line with NHS Digital guidance.
NHS Commissioning Board	NHS England (NHSE)
NHS Digital	Health Education England, NHS Digital and NHS England have merged into a single organisation. NHS England has assumed responsibility for all activities previously undertaken by NHS Digital.
NHS E	NHS England
NHS GM	NHS Greater Manchester Integrated Care Board
pseudonymised data	The process of distinguishing individuals in a dataset by using a unique identifier which does not reveal their 'real world' identity.
OPEL	Operational Pressures Escalation Levels – for further information see link here .

Secure Data Environments (SDEs)	Secure data environments are data storage and access platforms, which uphold the highest standards of privacy and security of NHS health and social care data when used for research and analysis. They allow approved users to access and analyse data without the data leaving the environment.
SDE's for R&D	Secure Data Environments accessed for health research were previously referred to by NHSE as "Trusted Research Environments" but have since been renamed "Secure Data Environments for Research & Development (SDE's for R&D)".
Snowflake	A cloud warehousing solution
Secondary Uses Service (SUS)	The Secondary Uses Service (SUS) is the single, comprehensive repository for healthcare data in England which enables a range of reporting and analyses to support the NHS in the delivery of healthcare services. When a patient or service user is treated or cared for, information is collected which supports their treatment. This information is also useful to commissioners and providers of NHS-funded care for 'secondary' purposes - purposes other than direct or 'primary' clinical care - such as: <ul style="list-style-type: none"> • healthcare planning • commissioning of services • National Tariff reimbursement • development of national policy
Trusted Research Environments (TREs)	See SDE's for R&D
Users	'Approved' individuals/business analysts that are supporting population health for the benefit of Greater Manchester
VCSE	The voluntary, community and social enterprise (VCSE) sector is an important partner for statutory health and social care agencies and plays a key role in improving health, well-being and care outcomes.

Section 1: Project Information

<p>Description, purpose of and reason for the initiative (GDPR Art. 35(7)): <i>Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of ...]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.</i></p> <p>1.1 Description, purpose and benefits:</p> <p>The Integrated Care Systems Design Framework and the "What Good Looks Like" framework³ articulate an expectation that Integrated Care Systems (ICSs) will have in place a linked, longitudinal dataset across primary and secondary care by March 2023 to enable population segmentation, risk stratification and population health management, expanding to other services (including social care) by 2024.</p> <p>The Goldacre Review ("Better, Broader, Safer: using health data for research and analysis") published in April 2022 recommended a shift from a model of data sharing and dissemination to a focus on data access, using "Secure Data Environments" to build public confidence and rigorously protect sensitive patient information (which has since been recognised as official DHSC policy).</p> <p>This is a rapidly evolving policy landscape, however Greater Manchester has mature digital and data assets in place which already support the delivery of each of these national requirements.</p> <p><u>Secure Data Environments (SDE's)</u></p>
--

³ [Success measure 7 - Healthy populations](#)

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): *Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.*

Historically, much of the analysis conducted within health & care services has operated based on a data sharing / data dissemination model, where data is extracted from one storage location and transferred to another. It is now recognised national policy that NHS data should shift towards a model of "data access" through the introduction of "Secure Data Environments" (SDE's).

Secure data environments are data storage and access platforms, which uphold the highest standards of privacy and security of NHS health and social care data when used for research and analysis. According to NHSE, SDE's provide an environment within which all health data should be accessed, therefore granting organisations a higher degree of control over who can become a user to view and interact with the data; what data they can see whilst in the environment; what users can do whilst in the SDE; and the information they can import or remove. This policy applies equally to all users of NHS data, who may be health and care analysts (operating at system or place level) or researchers.

SDE's are designed around the Office for National Statistics "[Five Safes Framework](#)", which are recognised as best practice in data protection

- safe settings - the environment prevents inappropriate access, or misuse
- safe data - information is protected and is treated to protect confidentiality
- safe people - individuals accessing the data are trained, and authorised, to use it appropriately
- safe projects - research projects are approved by data owners for the public good
- safe outputs - summarised data taken away is checked to make sure it protects privacy

Sub National secure data environments

NHS England announced over £13.5 million investment for teams across England to develop a country-wide network of NHS owned SDEs that will support the development of an interoperable network of NHS-owned Sub National Secure Data Environments, with further investment planned over 23/24 and 24/25.

The Sub National (SN) SDEs are NHS-led and bring together Integrated Care Boards with local universities and industry partners to build on existing collaborations and successful research partnerships.

Funding will ensure Sub National SDE coverage for the whole of England and was awarded to teams representing East of England; East Midlands; Great Western; Kent and Medway and Sussex; London; North East and North Cumbria; **North West**; Thames Valley and Surrey; Wessex; West Midlands; Yorkshire and Humber.

Sub national SDEs for research will offer near-real time, privacy protecting, access to rich linked data spanning different types including imaging, pathology and genomics. They will operate at significant scale, covering around 5 million citizens each, whilst preserving connectivity to local communities and clinical teams. The Sub National SDEs will be designed to operate smoothly with the NHS Digital (national) secure data environment, unified by a community of practice that will inform SDE policy and build on prior investments including – where lawful – the use of information from shared care record solutions.

The North-West SN SDE includes:

- Cheshire and Merseyside
- Greater Manchester
- Lancashire and Cumbria

SDE's for Research & Development

Alongside users within the health & care system, SDEs also support access to health data for research purposes, for both academic and commercial organisations. Providing researchers with access to high-quality data assets is a key priority as recognised by several national policy documents, including the [Life Sciences Vision](#), the [Future of Clinical Research Delivery](#), and the [Levelling Up White Paper](#), and more recently the Hewitt review of integrated care systems supported by [substantial investment \(£200m\) from NHSE and BEIS](#).

Secure Data Environments accessed for this purpose were previously referred to by NHSE as "Trusted Research Environments" but have since been renamed "Secure Data Environments for Research & Development (SDE's for R&D).

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): *Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.*

Regional context: GM Secure data environment – the Analytics & Data Science Platform (ADSP)

The GM SDE is a component of the GM Analytics & Data Science Platform (GM ADSP). The ADSP is a combination of "best of breed" technologies intended to serve GM ICS partners including NHS GM ICB,, local authorities, health providers, GMCA, NHSE and VCSE users. The ADSP, contracted by NHS GM, is hosted by the DSCRO for Greater Manchester, Arden & Gem Commissioning Support Unit (AGEM CSU).

The ADSP will provide GM's longitudinal record to support direct care and non-direct care uses. In the first instance this will be by securely linking data from primary care (GP records from the GMCR) and secondary care (from SUS) at patient level, with the longitudinal record expanding over time as additional data are linked. The ADSP is GM's recognised platform for conducting all forms of analysis and research – including PHM, risk stratification, planning and evaluation – and the functionality will be continually enhanced with input from users across the city-region. The medium- term strategy for the continuous development of ADSP, and how it aligns to national requirements and local need, is articulated in the [GM Analytics Prospectus](#), including GM's ambition to develop direct care dashboards that will enable the appropriate re-identification of patients identified where there is a clear clinical justification for doing so, and a legal basis (such as complex case finding for patients at-risk).

In addition, SDEs empower the accelerated adoption of innovation by providing the capability to conduct large-scale, population-level research generating quality outcomes data surrounding a new intervention, evaluating its impact in an uncontrolled environment, and facilitating the translation of any insights generated into a live operational setting as part of a learning health system.

In this way, the SDE itself and the underpinning workstreams form an integral part of the infrastructure for the recently announced Innovation Accelerator projects, to provide GM-wide shared data infrastructure for evaluation and insight as well as creating valuable, re-useable data assets.

Building upon recognised routes for the processing of NHSE data into AGEM CSU, the ADSP combines standard commissioning datasets with local data flows. It contains several components including Snowflake, a cloud warehousing solution, and GM-wide Tableau for disseminating real-time data visualisation. Tableau, is NHS GM's recognised tool for business intelligence, has been widely adopted at both GM and locality level providing a single source of the truth for analytics and collaborative working across localities.

However, the ADSP has been designed to be agnostic to the analytical tools a team might want to use, whilst also providing a core suite of solutions that all GM partners have equal access to. In addition to the primary use of the ADSP to provide descriptive and reporting insight to GM partners, the ADSP has been designed to maximise our capabilities regarding predictive modelling, population health management (PHM) and advanced analytical techniques including data science and machine learning.

The GM SDE forms part of a wider regional collaboration – the North-West SDE.

See also Appendix E for more information on the secure data environment guidelines.

GM Governance

The ADSP fundamentally sources data from 3 sources:

1. National datasets collected by NHSE. These datasets are shared with NHS GM ICB under the terms of a contract and Data Sharing Agreement between NHSE and the ICB. The NHSE Sub-License approach allows the ICB to enter into data sharing agreements with its ICS partners to share data these datasets. The data shared will be limited to pseudonymised data without the providers unique local patient id included.
2. Local data flows sent directly by health and care providers e.g., hospitals or local authorities.
3. Data flows extracted from our local shared care record (GM Care Record)

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): *Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.*

Applying to use data in the GM SDE

Three governance groups manage access and use of the ADSP.

1. SDE Data Access Committee (DAC). This group is responsible for approving applications to use the data held within the SDE for direct care, secondary use, and research purposes. This group considers the merit and ethnics of the applications. The group is chaired by the NHS GM Caldicott Guardian and is attended by public members, data controller representatives, NHS GM Data Protection Office and NHS GM SIRO.
2. SDE Application Review Group. This is technical group consisting of Information Governance and Data Engineer expertise. It supports applicants with their SDE access requests, providing technical advice and makes recommendations to the DAC. All secondary use and research applications must be approved by this group prior to being considered by the DAC. The group is chaired by the NHS GM Deputy Director of Data and Analytics.
3. NHS GM ICB Information Governance Operations Group (NHS GM IGOG). This group may be required to sign off key documents pertaining to SDE applications e.g. Data Protection Impact Assessments and Sub Licencing Agreements. The group is chaired by the NHS GM SIRO and is attended by NHS GM DPO's and IT Security Lead.

1.2 How will you collect the data?

- National data sets from NHS England (for example Secondary Uses Services – SUS) will be provided to NHS GM ICB by North West DSCRO (operated by AGEM CSU)
- GM Care Record – data extracted directly by NHS GM
- Local data flows sent/collected directly from health and care providers by NHS GM ICB

Methods of data submission or capture are specified within data flows section in Appendix C

1.3 How will you use the data?

The data will support locality, multi-locality and GM wide analytics and reporting requirements as follows:

- Direct care e.g., case finding/clinical dashboards.
- Population health
 - Analysis of outcomes following certain health interventions (i.e., public health interventions as well as treatments)
 - risk stratification
 - cohort identification
 - health promotion
- Health and Care commissioning intelligence
 - capacity and demand planning
- GM governance – usage audits, data quality and system health reporting
- Research

1.4 Where and for how long will the data be stored? The data will be held in the GM ADSP as described in 1.1. and in Section 3. for the duration of the relevant contracts. See Appendix C for more information.

1.5 What processes will be in place to delete the data when it is no longer required to be retained?

The data will be destroyed in line with DSCRO and NHS GM data destruction protocols.

1.6 What is the source of the data? E.g., the individual themselves, 3rd party The GMCR data and NHS England (SUS etc. data). Other data sources will be added over time as set out in 1.2.

1.7 What will the data be used for? Clarify if the data will also be shared with anyone.

The data will be used to support the following activities:

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): *Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.*

- Direct care e.g., case finding/clinical dashboards.
- Population health
 - Analysis of outcomes following certain health interventions (i.e., public health interventions as well as treatments)
 - risk stratification
 - cohort identification
 - health promotion
- Health and Care commissioning intelligence
 - capacity and demand planning
- GM governance – usage audits, data quality and system health reporting
- Health and Care Research

1.8 Specify the demographic/cohort/criteria:

All patients registered with a NHS GM GP and individuals receiving treatment and care within GM

1.9 Specify the borough(s) or GM wide: GM wide

1.10 Specify the organisations involved in the processing (include any suppliers of e.g., databases):

Data controllers

- All data Controllers feeding data into the GMCR.
- NHSE for national and local commissioning datasets covered by the NHSE / ICB DSA
- NHS GM ICB for NHSE supplied data covered by NHSE / ICB DSA
- Health and Care Providers supplying local data flows direct to NHS GM ICB

Data Processors:

Snowflake Computing UK Limited – supplier of a data platform built for the cloud. Snowflake delivers the performance, concurrency and simplicity needed to store and analyse the data in one location. Snowflake is a solution for data warehousing, data lakes, data engineering, data science, data application development, and securely sharing and consuming shared data. See Appendix A for relevant data processing contract terms.

AGEM CSU (host MS Azure) – data management and DSCRO services

Sub-processors

For AGEM CSU:

- Microsoft - for secure Cloud data storage, infrastructure and support with Microsoft's Azure Cloud service.
- Greenworld Technologies, Stoke on Trent – for secure data and IT asset destruction.
- Trustmarque are supporting the CSU with using Microsoft's infrastructure. Trustmarque provide access, architecture, and design to assist organisations with their integration to Microsoft Cloud services. They also provide support for continuous efficiency of cloud services, to assist with resources and costings.

For Snowflake Computing UK Limited:

- MS Azure Cloud

Other 3rd parties

Interworks (Curator, and Matillion) – no processing of personal data.

Interworks (Tableau Cloud) using AWS – processing of pseudonymised patient level data with no access to pseudo keys.

University of Manchester (eLabs) - (provision of SDE for R&D) – no processing of personal data

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.

1.11 What contractual arrangements are in place (specify contract terms or embed or attach relevant sections of contract/SLA?)

NHS GM ICB has the contracts with the following:

Product	Parties to the contract	Date	Review / expiry date	Third party IP	Contract
Snowflake	1. Interworks 2. NHS GM	1/01/01/2025	31/12/2025	Snowflake Azure	G-cloud 13 call off contract
Matillion	1. Interworks 2. NHS GM	1/4/2023	31/3/26	n/a	G-cloud 13 call off contract
Tableau	1. Interworks 2. NHS GM	19/7/2023	18/7/2026	Salesforce	G-cloud 13 – no personal data processed This agreement will be replaced to reflect Tableau Cloud subject to completion of this DPIA
Curator	1. Interworks 2. NHS GM	30/05/22	31/05/25	n/a	G-cloud 12 – call off
Elabs	1. University of Manchester 2. NHS GM	22/03/2022	01/04/25	n/a	NHS Terms and Conditions – Schedule 3 covers Information and Data provisions
Arden & GEM	1. AGEM CSU 2. NHS GM	1/4/23	31/3/24	Azure	NHS Service Level Agreement New contract pending signature subject to this DPIA being completed.

In addition, the following data sharing agreements exist:

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.

National and local datasets supplied by NHSE:

- A data sharing framework contract between NHSE and NHS GM ICB
 - associated data sharing agreements with NHS England
- A data processing agreement with NHS Arden and Greater East Midlands Commissioning Support Unit (AGEM CSU) - (AGEM CSU is hosted by NHS England) - this is currently in development – see risk section

GMCR Data:

- Joint Controller agreement with all the GM Data Controllers allowing for the extraction of the data from the GMCR and associated DPIAs allowing for the transfer into the ADSP.

Local data flows supplied directly to NHS GM ICB by health and care providers.

- These will be subject to additional data sharing agreements as required.

1.12 How often will you be collecting and using the personal data? Daily

1.13 How long do you expect this initiative to last?

- ☐ | End of contract period
- ☐ | Specific time period – specify? [\[Click here to enter text\]](#)
- ☐ | Lifetime of system (where the initiative or project relates to a new or revised ICT system)
- ☒ | Other – specify In line with national requirements for SDEs with supporting contractual arrangements in place

1.14 What is the nature of your relationship with the individual data subjects for this initiative? This enables IG to ascertain the lawful basis for processing

Provision of health/social care ☒ | Protecting the health of the general public ☒ |

Local audit to assure safe health and social care ☒ | Checking quality of care, beyond local audit ☒ |

Supporting research ☒ | Staff employment ☐ | Other - specify: [\[Click here to enter text\]](#)

1.15 How much control will the data subjects have over the data being processed?

Patients can:

- Register an objection to having a shared care record via their GP practice, where, if upheld by the GP, a code can be applied to the GP Electronic Patient Record (EPR) to prevent the shared care record being created (including for direct care).
- Register an opt out via the National Data Opt out service to prevent their identifiable data being used for secondary use and research.
- Register a type one opt out (to prevent information being shared outside a GP practice for purposes other than direct care)
- Register a local opt out as follows
 - If a patient does not want their anonymised data from the GMCR to be used for planning and research within Greater Manchester, but do not want to apply a national opt out, they can contact us in one of the following ways:
 - By calling us on 0161 947 0770 and selecting option 8
 - Emailing us at contactus.caregateway@nhs.net
 - Or writing to us at:
Information Governance Team
NHS Greater Manchester Integrated Care
Tootal Buildings
Broadhurst House, 56 Oxford St,
Manchester M1 6EU

Description, purpose of and reason for the initiative (GDPR Art. 35(7)): Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.

(Nb: the application of local opt outs for secure data environments is currently under review by NHS England in discussion with the Confidentiality Advisory Group. A local opt out will be applied to the GM SDE until we are advised by NHS England and the Confidentiality Advisory Group that a local opt out is not required.)

See also Appendices B1 and B2.

1.16 Would they expect you to use their data in this way?

Yes ☒ | No ☐ | Don't know ☐

[Click here to enter text.](#)

The following gives national examples of public views on using data for planning and research:

- [BMJ - Public opinion on sharing data from health services for clinical and research purposes without explicit consent: an anonymous online survey in the UK](#) – first published 27 April 2022
- [Putting Good into Practice: a public dialogue on making public benefit assessments when using health and care data](#) (co funded by the National Data Guardian for Health and Social Care, Understanding Patient Data and UK Research and Innovation's Sciencewise programme – April 2021
- [The Lancet – Digital Health - Public perceptions on data sharing: key insights from the UK and the USA – Published July 2020](#)
- [HRA/University of Sheffield - Public views on sharing anonymised patient-level data where there is a mixed public and private benefit - September 2019](#)
- [Understanding Patient Data: Public attitudes to patient data use – July 2018](#)
- [NHS Health Research Authority – Survey of the general public: attitudes towards health research](#) – 2013

1.17 How will you consult with them to seek their views on the data processing – or justify why it is not appropriate to do so:

A significant amount of public engagement work has been undertaken in relation to the GMCR and the use of patient data for secondary use and research.

Public engagement activity is planned for the ADSP, and this work will build on the existing links we have in seldom heard communities in Greater Manchester that were engaged with as part of the communications campaign for the GM Care Record. These existing groups (including the black African/Caribbean communities, South Asian communities, communities with high rates of deprivation and older age groups) will be revisited as part of the work on the SDE and the use of data without consent.

In addition, Health Innovation Manchester, The University of Manchester and our other health and care partners have existing patient and public engagement groups that we will take this issue with to get their expert advice and views on.

Public posters have been disseminated to all data controllers explaining the implementation of a local patient opt out and the public facing website has been updated accordingly.

There is public representation on the Secondary Uses and Research Group where applications for data from the GM Care records are reviewed/approved. A specific website has been set up to support broader transparency about the GMCR data use <https://gmwearebettertogether.com>

1.18 Do you need to consult with anyone else internally or externally?

Consultation will be undertaken via applicable GM governance

1.19 Will individual's personal information be disclosed outside of the parties to this initiative in identifiable form and if so to who, how and why?

☐ | Yes – provide details below ☒ | No ☒

[Click here to enter text](#)

1.20 If the information is to be anonymised or pseudonymised in any way, specify how this will happen

See Appendix C item 2.

<p>Description, purpose of and reason for the initiative (GDPR Art. 35(7)): Specify how many individuals will be affected or state the detail in relation to the demographic e.g., all adults over the age of 65 in the [area/borough(s) of]. Embed any relevant project documentation e.g., PID, service specification, business case, flow diagrams of how the data will be processed.</p>
<p>1.21 If personal data is being transferred outside of the EEA, describe how the data will be adequately protected (e.g. the recipient is in a country which is listed on the Information Commissioner's list of "approved" countries - see link here). (This would include database/information hosted on ICT applications outside the UK)</p> <p>No data is transferred outside the EEA, data within the SDE is stored and processed within the UK.</p> <p>Snowflake being a US provider provide support services globally. Whilst the data always remains hosted in the UK and it expects to carry out most support activities out of its UK operations, it may also have to use its affiliates in other jurisdictions including the US, Singapore, Brazil, i.e., countries which are not on the list of "approved countries".</p> <p>The agreement agreed with Interworks, Snowflake's distributor, incorporates the EU model clauses together with the UK addendum to ensure compliance with Article 47 of the GDPR.</p> <p>All other processors and sub-processor process the data in the UK only.</p> <p>1.22 Are there any approved national codes of conduct or sector specific guidelines that apply to the data e.g., ICO/DoH&SC/NHS England/NHS Digital etc. (GDPR Art. 35(8)) (Remove or add to the below list as necessary)</p> <ul style="list-style-type: none"> • Secure data environment for NHS health and social care data – policy guidelines • Sub licensing guidance – v6 • GOV.UK NHS Constitution – updated Jan 2021 • GOV.UK Handbook to the NHS Constitution – updated Feb 2021 • Codes of practice for handling information in health and care • BMA guiding principles – Disclosing patient data for secondary purposes – updated September 2020 • ICO - Anonymisation: managing data protection risk code of practice • NHS Digital - ISB1523: Anonymisation Standard for Publishing Health and Social Care Data • National Data Guardian (NDG) guidance - What do we mean by public benefit? Evaluating public benefit when health and adult social care data is used for purposes beyond individual care • Enisa – European Union Agency for Cyber Security - Pseudonymisation techniques and best practices - Recommendations on shaping technology according to data protection and privacy provisions - NOVEMBER 2019 • NHS Digital Clinical Information Standards • HM Government's Technology Code of Practice • UK Government's Open Standards Principles • NHS Digital, Data and Technology Standards • Department of Health Social Care Code of Conduct for data-driven health and care technology dated February 2019 • NHS Digital Clinical Risk Management Standards – DCB0129 and DCB0160 • Information Governance Framework for Integrated Health and Care: Shared Care Records – published September 2021 <p>1.23 How will you prevent function creep i.e., the gradual widening of the use of a technology or system beyond the purpose for which it was originally intended, especially when this leads to potential invasion of privacy? Access to and uses of the ADSP (as set out in this DPIA) will be monitored via GM ICS Governance Groups and NHS GM Data Governance Lead.</p> <p>1.24 How will you ensure data quality? Source providers are responsible for their own data feeding into the GMCR and into SUS however we will develop a set of systematic data quality routines (see risk section).</p>

Section 2: Data Items

An electronic data catalogue containing all the national and local data flows is currently being developed by NHS GM to document all the data sets and data items flowing into the ADSP.

DPIA NHS GM Secure Data Environment – Analytics and Data Science Platform (ADSP) V1.0

The data catalogue will consist of datasets within the SDE and those embedded within Tableau.

This will be available to all end users of the ADSP to support analysis.

Section 3 – Data Flows – See Appendix C – Item 3.

Section 4: Information Technology -

List any applicable electronic systems/software to this initiative (current and/or new):

4a) System name	Used by e.g., organisation and dept.	Parties/system supplier
Azure	AGEM CSU / DSCRO Tenancy	AGEM MS Azure
Snowflake	NHS GM – BI Team	Interworks reseller for Snowflake
Tableau	NHS GM – BI Team	Interworks reseller for Tableau
E-Labs	NHS GM – BI Team	University of Manchester TRE
Curator	NHS GM – BI Team	Interworks Curator
Matillion	NHS GM – BI Team	Interworks reseller for Matillion
Patient Re-ID application	NHS GM – BI Team	AGEM CSU
Patient Re-ID User Management system	NHS GM – BI Team	NHS GM in-house development
Amazon Web Services	NHS GM – BI Team	Interworks reseller for Tableau Cloud which uses the AWS cloud platform


4c)

4b) Confirmation of IT involvement – IT lead(s)/support

Name	Organisation	Involved Y/N but planned
Graham Beales	NHS GM	Y

other assets: Specify any other relevant assets relating to the personal data being processed either in use or intended

Asset name e.g., child health record	Format e.g., paper/excel spreadsheet	Asset id (linked to organisation information asset register) – if not yet registered leave blank
Not applicable		

4d)	Where a data system is in use as part of the project/initiative confirm the following:	
i)	<p>Appropriate technical & organisational security measures in place to protect data.</p> <p>(Including specifications, information security policies, certifications (e.g., ISO27001), independent penetration test reports for any application/database and hosting infrastructure)</p> <p>where cloud computing is being utilised ensure sufficient security in place as in attached appendix within the contract or complete the embedded document and attach as an Appendix.</p>  <p>Appendix%20-%20C loud%20data%20stc</p>	<p>Yes <input checked="" type="checkbox"/> Explain process or attach relevant documentation:</p> <p>See Appendix C for data flows specifying security controls for each flow and Appendix D for security measures in place regarding Snowflake.</p> <p>No <input type="checkbox"/> If no, explain: Click or tap here to enter text.</p>
ii)	<p>Staff access is audited.</p>	<p>Yes <input type="checkbox"/> Explain process: The NHS GM ICB Data Services Team monitor user access logs to the components within the GM ADSP and we are in the process of developing audit reports for the NHS GM ICB Data Governance Lead – risk GMSDE10</p> <p>No <input type="checkbox"/> If no, explain: Click here to enter text.</p>
iii)	<p>Appropriate role-based access controls are in place for all staff who have access:</p>	<p>Yes <input checked="" type="checkbox"/></p> <p>See Appendix C item 5, and Appendix F</p> <p>No <input type="checkbox"/> If no, explain: Click here to enter text.</p>

iv)	A Digital Technology Assessment Criteria been completed. <i>The DTAC is a non-mandatory assessment tool which provides assurance to NHS organisations that suppliers are meeting required standards..</i> NHSE DTAC Link	Yes <input type="checkbox"/> No <input type="checkbox"/> Not applicable ✓
v)	An Information Asset Owner (IAO) and Information Asset Administrator (IAA) been assigned for the system	Yes (specify below) <input checked="" type="checkbox"/> No <input type="checkbox"/> Don't know <input type="checkbox"/> IAO: Chief Analytics and Intelligence Officer, NHS GM IAA: Data Warehouse and Infrastructure Manager, NHS GM

Section 5: Information governance project assurance (to be completed by Information Governance)

GDPR Article 35(3) and ICO guidance 35(4)		Yes	No	Unsure	Comments <i>Document initial comments on the issue and the privacy impacts or clarification why it is not an issue</i>
i)	Is there to be systematic and extensive profiling with significant effects : "(a) any systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person"	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Where profiling takes place there will be clinical intervention so that the processing is not solely automated.
ii)	Is there large-scale use of sensitive data : "(b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10"	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	GMCR patient data and SUS data plus other local data sources
iii)	Is there monitoring of the public : "(c) a systematic monitoring of a publicly accessible area on a large scale"	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	[Click here to enter text]
iv)	Does the processing involve the use of new technologies , or the novel application of existing technologies (including AI).	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Click here to enter text.
v)	Is there any denial of service : Decisions about an individual's access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Click here to enter text.
vi)	Does the initiative involve profiling of individuals on a large scale ?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Risk stratification and population segmentation
vii)	Is there any processing of biometric data?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Click here to enter text.
viii)	Is there any processing of genetic data other than that processed by an individual GP or health professional, for the provision of health care direct to the data subject?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	[Click here to enter text]
ix)	Is there any data matching : combining, comparing or matching personal data obtained from multiple sources?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Click here to enter text.
x)	Is there any 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 Article 14 would prove impossible or involve disproportionate effort.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	[Click here to enter text]
xi)	Is there any tracking of individuals: processing which involves tracking an individual's geolocation or behaviour, including but not limited to the online environment.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	[Click here to enter text]

GDPR Article 35(3) and ICO guidance 35(4)		Yes	No	Unsure	Comments <i>Document initial comments on the issue and the privacy impacts or clarification why it is not an issue</i>
xii)	Is there any targeting of children or other vulnerable individuals : The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Profiling as referenced above.
xiii)	Is there any risk of physical harm : Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Patient level data for all GM GP registered patients

				Action required – ensure covered in section 6
5.1	Is the initiative supporting the delivery of direct care ⁹ ?	Yes <input checked="" type="checkbox"/>	No <input type="checkbox"/>	
5.2	Is it supporting the delivery of any other main purpose?	No <input type="checkbox"/> Yes <input checked="" type="checkbox"/> Commissioning <input checked="" type="checkbox"/> Public health <input checked="" type="checkbox"/> Monitoring health and social care <input checked="" type="checkbox"/> Research <input checked="" type="checkbox"/> Related to staff employment <input type="checkbox"/> other <input type="checkbox"/> specify: [Click here to enter text]		
5.3	Are the arrangements for individual's to either object to their information being shared for direct care or to opt-out of the initiative for indirect care, once they have been provided with appropriate communication about it, appropriate? (See 1.4 – 1.6)	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>Specify any action required and document in action plan at section 6` Ongoing review and implementation of SDE communication and engagement plan	
5.4	Confirm appropriate subject access handling/information rights procedures in place?	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>	state reason if no - [Click here to enter text] NHS GM SARs process	
5.5	Who are the controllers in this initiative?	Not applicable <input type="checkbox"/> GM Health and Care Providers, Practices, NHS GM and NHS England (SUS data)		
5.6	Are there any data processors and have the processors had oversight and opportunity to input into this DPIA?	Not applicable – no processors <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Planned <input type="checkbox"/> Don't know <input type="checkbox"/>	Processors are listed in section 1.10 and have had opportunity to contribute to this DPIA	
5.7	Are the contractual terms at 1.11 sufficient to satisfy IG?	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Don't know <input type="checkbox"/>	The contracts with Interworks and AGEM have been reviewed legally by the legal consultancy, V-lex Limited. The relevant extracts of the Snowflake contract are attached in Appendix A.	
5.8	Does each party confirm that information governance training is in place and all staff with access to personal data have had up to date training	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Don't know <input type="checkbox"/>	NHS GM and AGEM CSU – Mandated as part of DSPT Snowflake – Confirmed	

⁴ The definition of direct care is A clinical, social or public health activity concerned with the prevention, investigation and treatment of illness and the alleviation of suffering of individuals. It includes:-

- supporting individuals' ability to function and improve their participation in life and society
- the assurance of safe and high-quality care and treatment through local audit,
- the management of untoward or adverse incidents
- person satisfaction including measurement of outcomes

undertaken by one or more registered and regulated health or social care professionals and their team with whom the individual has a legitimate relationship for their care

			Action required – ensure covered in section 6
5.9	Confirm all parties have appropriate measures in place to report incidents and share learning?	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Don't know <input type="checkbox"/>	AWS – Confirmed via DSPT standards exceeded NHS GM and AGEM CSU – Mandated as part of DSPT Snowflake – confirmed NHS GM – confirmed AWS – Confirmed via DSPT standards exceeded
5.10	Does each party involved in the processing of NHS personal identifiable data complete a Data Protection and Security Toolkit Assessment or undertake another recognised standard?	NHS GM and AGEM CSU complete the DSPT assessment Snowflake – Standards Exceeded AWS Standards Exceeded	See risk section re Snowflake
5.1 1	Has each party involved in the processing paid the ICO registration fee? https://ico.org.uk/about-the-ico/what-we-do/register-of-fee-payers/	Yes <input checked="" type="checkbox"/>	NHS GM – ZB343633 – renewal 28 June 2025 AGEM – Z2950066 renewal date 17 Feb 2025 Snowflake – ZA763210 renewal date 23 June 2025 – Amazon Web Services EMEA SARL - ZA481902 renewal 12th December 2024
5.12	Does there need to be an Information Sharing agreement between the relevant parties that covers the processing arrangements?	Yes <input checked="" type="checkbox"/> – specify reasons why: Whilst NHS GM is the only controller for the ADSP, it is receiving data from other databases. These data flows need to be covered by data sharing agreements. All data extractions from the GMCR are covered by the existing joint controller agreements except for adult social care. Data Sharing Framework Contract and DSA exist between NHS GM and NHSE for national and local data flows Data Sharing Agreements also exist between NHS GM and Local Authorities for Adult Social Care data flowing directly into the SDE Any new extractions or uses not previously covered may need a separate DPIA and DSA	
5.13	Confirm all relevant organisations have appropriate cyber security measures and/or are working towards cyber essentials	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Don't know <input type="checkbox"/> Attach or embed confirmation e.g., email from IT if yes:	AGEM – Cyber Essentials Plus & ISO 9001 and 14001 Snowflake – as part of the contract terms Snowflake agrees to obtain the Cyber Essentials + notification as soon as reasonably practicable. They are also ISO27001, SOC 2 Type II and SOC 1 Type II certified / audited. Amazon Web Services – supports a wide range of security standards including Cyber Essentials. Please see the following link <u>United Kingdom (amazon.com)</u>

		Action required – ensure covered in section 6
5.14	Lawful Basis for processing:	
The Health and Social Care (Safety and Quality) Act 2015 inserted a legal Duty to Share Information in Part 9 of the Health and Social Care Act 2012 (health and adult social care services: information) https://www.legislation.gov.uk/ukpga/2015/28/pdfs/ukpga_20150028_en.pdf		Tick if applicable: <input type="checkbox"/>
Official Authority:		
GP Practices	NHS England's powers to commission health services under the NHS Act 2006.	<input checked="" type="checkbox"/>
NHS Trusts	National Health Service and Community Care Act 1990	<input checked="" type="checkbox"/>
NHS Foundation Trusts	Health and Social Care (Community Health and Standards) Act 2003	<input checked="" type="checkbox"/>
Local Authorities	Localism Act 2011 Children Act 1989 Children Act 2004 Care Act 2014	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>
GDPR	Article 6 condition(s) for processing: (e) Public task Choose an item. Choose an item.	Article 9 condition(s) for processing: (h) Health or social care (i) Public Health (j) Archiving, research and statistics
DPA 2018	Schedule 1, Part 1, condition(s) for processing: (2) Health or social care If health and care is selected specify the purpose below: (f) management of health care systems or services or social care systems or services If public health is selected, confirm the processing is carried out: (b)(i) by or under responsibility of a health professional (b)(ii) by another person who in the circumstances owes a duty of confidentiality under an enactment or rule of law If research is selected confirm the that the processing: (a) is necessary for archiving purposes, scientific or historical research purposes or statistical purposes (b) is carried out in accordance with Article 89(1) of the GDPR , and (c) is in the public interest Confirm <input checked="" type="checkbox"/>	
Human Rights Act	Legitimate aim For the following reason (if applicable) protect health or morals	
Common Law duty of Confidentiality	Statutory basis or legal duty to disclose The use the ADSP for secondary use and research has been approved by the Secretary of State for Health following successful application to the National Confidentiality Advisory Group	
National Data Opt out (The national data opt-out allows a patient to choose if they do not want their confidential patient information to be used for purposes beyond their individual care and treatment - for research and planning.) For more information see link here .		
To be applied where relevant		

Section 6 – Privacy issues identified and risk analysis

Consider the potential impact on individuals and any harm or damage that might be caused by your processing – whether physical, emotional, or material. In particular, look at whether the processing could possibly contribute to:

DPIA NHS GM Secure Data Environment – Analytics and Data Science Platform (ADSP) V1.0

- unauthorised access to data
- inability to exercise rights (including but not limited to privacy rights);
- undesired modification of data
- inability to access services or opportunities;
- disappearance of data
- discrimination;
- loss of control over the use of personal data;
- identity theft or fraud;
- reputational damage;
- financial loss;
- loss of confidentiality;
- physical harm;
- re-identification of pseudonymised data; or
- any other significant economic or social disadvantage

Include any sources of the risk i.e. person or non-human source that can cause a risk either accidentally or deliberately:

Source of risk	Examples
Internal human sources	A negligent or rogue employee, proximity of the system, skills, privileges and available time are potentially high, possible lack of training and awareness
External human sources	A rogue or naive neighbour, by having a physical proximity, hacking into the devices data
Non-human sources	Incident or damage at one of the organisations (power cut, fire, flood, etc.)

Specify any issues identified, recommendations and actions needed to secure the data if appropriate controls not in place within the risk assessment.

The risks should be reviewed, scored using the risk matrix below and incorporated into a risk register.

The level of risk is scored out of 25. A score of 0-5 is attributed to both the impact on the rights and freedoms of the individual, and the likelihood of those rights and freedoms being compromised. The two scores are then multiplied to create the composite risk score using the risk matrix below. This should be recalculated in the final columns to take into account proposed solutions/actions.

Risk	Description	Risk Score see matrix below			Proposed solutions/actions	Responsibility and date	Revised risk score when actions addressed see matrix below		
		Impact	Likelihood	Risk rating			Impact	Likelihood	Risk rating
GMSDE1 – unauthorised access to data	Insufficient governance in place to approve applications to use the ADSP	4	3	12	Undertake governance review identifying any gaps and implementing revised structure as necessary. This must include the sub-licensing arrangements.	NHS GM Head of BI, Manchester in collaboration with NHS GM SIRO and system partners Action completed: implemented Data Access Committee	3	1	3
GMSDE2 – unauthorised access to data	Application form(s) and processes insufficient to capture required information from applicants	3	4	12	Agree application form to ensure all data feeds are suitably captured	NHS GM Head of BI, Manchester in collaboration with NHS GM SIRO and system partners Action completed: application forms developed and implemented.	3	1	3
GMSDE3 – unauthorised access to data	User Management process is not robust e.g., vetting of users, removal of access when projects finish, or staff leave.	4	4	16	Review of existing user management processes, identify and address any weaknesses	NHS GM Deputy Director of Data and Analytics (Data Governance) collaboration with NHS GM SIRO July 2023 – March 2025	4	2	8

GMSDE4 – unauthorised access to data	1.11 – Insufficient contractual arrangements in place with data processors prior to processing of personal data	5	2	10	NHS GM to finalise and sign off contractual arrangements with data processors prior to permitting the processing of personal data	Awaiting signature of AGEM CSU SLA for 2024/25 Awaiting signature of contracts with Interworks for Snowflake and Tableau Cloud subject to sign off of this DPIA	5	1	5
GMSDE5 – undesired modification of data	1.24 – insufficient set of systematic data quality routines to ensure data is of suitable quality	3	5	15	Source providers are responsible for their own data feeding into Graphnet and into SUS however NHS GM will develop a set of systematic data quality routines	NHS GM Deputy Director of Data and Analytics NHS GM has developed a Data Quality Policy and processes. These need to be implemented in 2025.	5	2	10
GMSDE6 – data subjects not aware of access rights and ability to opt out	1.17 – Communication materials not in place – public website and posters	4	3	12	1. Update public website to cover SDE and use cases with ongoing reporting of data uses 2. Keep data controllers informed of processing and use cases 3. Disseminate posters for data controllers to display explaining right to opt out of use of data for planning and research	Head of Communications – GMCR/GM SDE July – December 2023 (with ongoing monitoring and updating of website) Action completed: wide scale public communications and engagement exercise undertaken	3	1	3
GMSDE7 – Unauthorised access to data	5.10 and 5.11 Data processor Snowflake need to confirm DSPT completed and ICO registration fee paid and visible on ICO register	4	3	12	1. Snowflake to confirm ICO registration prior to processing. 2. Snowflake to confirm DSPT submission prior to processing.	NHS GM Head of BI, Manchester and HInM Head of IG for SDE August 2023 Action completed	3	1	3
GMSDE8 – Unauthorised access to data	5.12 social care data – insufficient IG compliance in place to process social care data for secondary uses and research	5	2	10	1. Separate DPIA's covering Adult Social Care Include social care data as part of CAG application in collaboration with social care or seek amendment to NHS GM CAG application when social care IG compliance in place	NHS GM Deputy Director of Data and Analytics and HInM Head of IG for SDE 2025	5	1	5
GMSDE9 – unauthorised processing of data	5.14 Common Law Duty of Confidentiality to be satisfied to process data	5	3	15	Submit CAG application	NHS GM Head of BI, Manchester and HInM Head of IG for SDE Action completed: S251	3	1	3

						CAG application approved for secondary use and research			
GMSDE10 - unauthorised processing of data	Failure to monitor access to data within agreed parameters	3	4	12	Develop Standard Operating procedure for audit of access to data	Deputy Director of Data and Analytics 2025	3	1	3

	Impact (How bad it may be)		Likelihood (The chance it may occur)			Risk Rating Likelihood x Impact = TOTAL RISK RATING				
						Impact				
						1	2	3	4	5
5	Very High (Will have a major impact)	5	Almost certain (almost certain to happen/recur; possibly frequently)	Likelihood	5	5	10	15	20	25
4	Major (highly probable it will have a significant impact)	4	Likely (Will probably happen/recur, but is not a persisting issue or circumstance)		4	4	8	12	16	20
3	Moderate (Likely to have an impact)	3	Possible (Might happen or recur occasionally)		3	3	6	9	12	15
2	Minor (May have an impact)	2	Unlikely (Do not expect it to happen/recur, but it is possible it may do so)		2	2	4	6	8	10
1	Negligible (Unlikely to have any impact)	1	Rare (This probably will never happen/recur)		1	1	2	3	4	5

Total Risk Rating	Risk
1-3	Low
4-6	Moderate
8-12	High
15-25	Extreme

Section 7 – Conclusion (tick one of the following)

☒ All privacy risks have been identified and actions are underway to mitigate, accept or remove the risks. This action plan will now be reviewed and monitored via NHS GM Information Governance Operational Group

[\[Click here to enter text\]](#)

☐ All privacy risks have been identified and actions completed to mitigate, accept or remove the risks

☐ Not all privacy risks can be removed or reduced, and the processing remains high risk, therefore the ICO must be consulted

Nb. Where the processing remains high risk, that cannot be mitigated or remove, the ICO must be consulted:

ICO Review required Yes ☐ No ☒

If yes, ICO review outcome and date [\[Click here to enter text\]](#)

[Click here to enter a date.](#)

Section 8: Approval and Sign off (this can be configured to reflect local arrangement for sign off if required – some may want the DPO to sign off, others may not. However, the DPO should review all DPIAs)

Approved by:

Organisation	Name	Date
NHS GM Integrated Care	Information Governance Operational Group	24/08/2023

For [enter approval body] use only – Nb. The following can be completed by each organisation and retained locally – it does not need to be collated for each organisation involved

Data Protection officer (DPO) review	<input checked="" type="checkbox"/>	Name and organisation: NHS GM DPO – Shavamah Purves Click here to enter a date.
DPO review not required	<input type="checkbox"/>	Decision made by: Click here to enter text.
Approved – no actions required	<input type="checkbox"/>	Click here to enter a date.
Approved with action plan	<input checked="" type="checkbox"/>	31/10/2023
Declined (give reason)	<input type="checkbox"/>	Click here to enter text. Click here to enter a date.
Senior Information Risk Owner - signature		
Incorporate data flows into data flow mapping or onto the Information Sharing Gateway (ISG)	<input type="checkbox"/>	Click here to enter a date.
Incorporate assets into the asset register or onto the ISG	<input type="checkbox"/>	Click here to enter a date.
Confirm staff handling subject access requests are aware of new or changed information asset	Yes <input type="checkbox"/> Not applicable <input type="checkbox"/>	Click here to enter a date.
Confirm Information Sharing arrangements documented: <ul style="list-style-type: none"> within this DPIA and ISA not required <input type="checkbox"/> within a separate IS agreement <input checked="" type="checkbox"/> uploaded into the Information Sharing Gateway <input type="checkbox"/> planned within the DPIA action plan <input type="checkbox"/> Within a Data processing contract <input type="checkbox"/> Other: specify - Click here to enter text.		24/08/2023
Monitor and review of this DPIA	Who by: NHS GM IGOG	When Ongoing

Appendix A – Snowflake contractual terms



Appendix B

Greater Manchester Care Record (GMCR)

Patient notification

The Greater Manchester Care Record (GMCR) is used by health and care professionals to make sure you receive the best possible treatment and care. It pulls together information from several important areas of health and care within the Greater Manchester including:

- Primary care e.g., GP practices
- Community services
- Mental health services
- Social care
- Secondary care e.g., hospitals
- Specialist services e.g., North West Ambulance Service (NWAS) and the Christie

It includes:

- Health or care issues we need to be aware of
- Any medications you may be taking
- Any allergies you may have
- Recent test results
- Vaccinations
- Care or treatment plans
- Details of any social care or carer support you may receive

A record of care is held on each organisation's secure electronic system (local record) e.g., a GP practice will have their own system for recording patient information. Graphnet, a supplier of healthcare systems, has designed a secure system that integrates data from those multiple electronic health and social care systems to provide a live summary of that data to a health or social care worker when required for the purposes of your individual care.

The GM Care Record:

- Safely speeds up decisions related to your care
- Provides access to vital information such as medication, test results and allergies
- Reduces instances of harm, such as allergic reactions
- Enhances care planning
- Improves care coordination across all services and boroughs

DPIA NHS GM Secure Data Environment – Analytics and Data Science Platform (ADSP) V1.0

NHS Greater Manchester Integrated Care have also designed a secure analytics platform, where, working with technical staff at NHS England, they create a de-identified copy of that data by removing identifiers e.g., your name, address date of birth etc. This deidentified data can then be linked with other national and or local data sets and used to support important research and the planning of health and care services (non-research). Whilst you are automatically enrolled into the GM Care Record as a GM citizen, you have the option to object to your information being shared for individual care and to opt out of your data being used for research and planning.

We apply the Type 1 opt out⁵ implemented via your GP practice and the National Data Opt out⁶ already via the secure analytics platform. If you have either of these opt outs applied to your record your GMCR data will not flow into the secure platform for planning and research.

However, if you do not want to apply a national opt out and just opt out of the GM Care Record data being used for planning or research within Greater Manchester you can contact us as follows:

- Calling us on 0161 947 0770 and select option 8
- Emailing us at contactus.caregateway@nhs.net
- Or writing to us at:

Information Governance Team
NHS Greater Manchester Integrated Care
The Tootal Buildings
56 Oxford St
Manchester M1 6EU

The Health Research Authority (HRA) (for research activity) and the Secretary of State for Health and Social Care (for non-research) has given Section 251 support for the activity following advice from the Confidentiality Advisory Group.

The CAG approval reference is as follows:

23/CAG/**7 – Non research

23/CAG/** – Research

For further information on the GM Care Record speak to your care provider or access the following link: <https://gmwearebettertogether.com/>

If you require information in an alternative format, please contact us via one of the methods listed above.

⁵ <https://www.nhs.uk/using-the-nhs/about-the-nhs/opt-out-of-sharing-your-health-records/>

⁶ <https://digital.nhs.uk/services/national-data-opt-out>

⁷ *** to be added CAG application reference is received

Appendix C

Analytics and Data Science Platform (ADSP) - Infrastructure and data flows

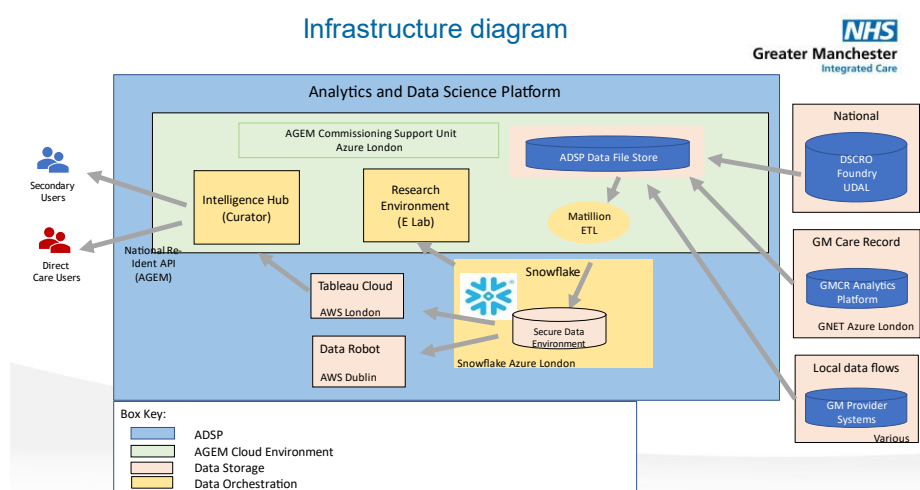
Introduction

This appendix describes the following:

1. The ADSP infrastructure – components and data storage locations
2. The pseudonymisation process and a summary of the governance arrangements
3. Data flows into the Secure Data Environment (SDE)
4. Summarises the User Management process
5. Role Based Access Control Groups for the SDE

1. Infrastructure diagram

Commented [A1]: Updated to reflect Tableau Cloud and use of AWS London



2. Summary of the pseudonymisation and governance process

1. NHS GM ICB and the DSCRO operate a 2 stage pseudonymisation process.
2. The DSCRO and the ICB make use of the Open Pseudonymiser or the DSCRO's PIPE Tool to pseudonymise the data.
3. DSCRO will manage the pseudonymisation SALT (encryption) keys.
4. All data will be pseudonymised, even data used to support direct care.
- 5.
6. All direct care use cases will make use of the NW DSCRO Patient Re-ID application.
7. NHS GM ICB will establish governance arrangements as per NHSE ICS Sub Licence guidance to approve patient re-id use cases.
8. All local patient identifiable datasets flowing directly into the ADSP will be subject to appropriate data processing / sharing agreements.

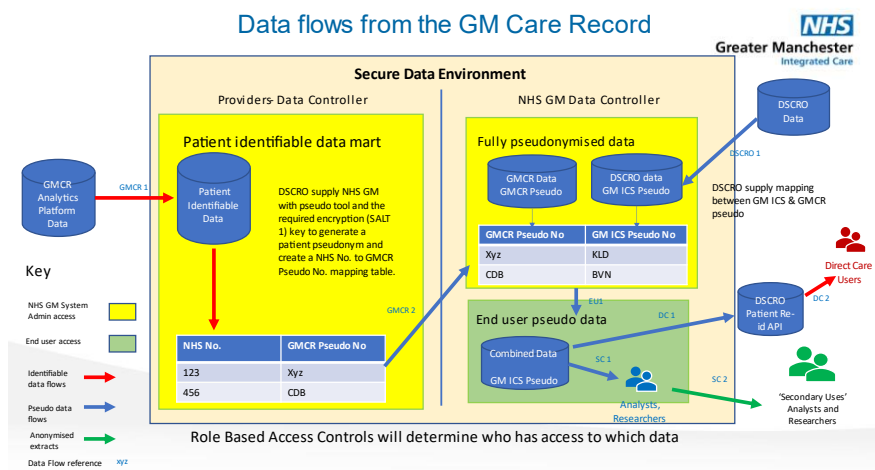
DPIA NHS GM Secure Data Environment – Analytics and Data Science Platform (ADSP) V1.0

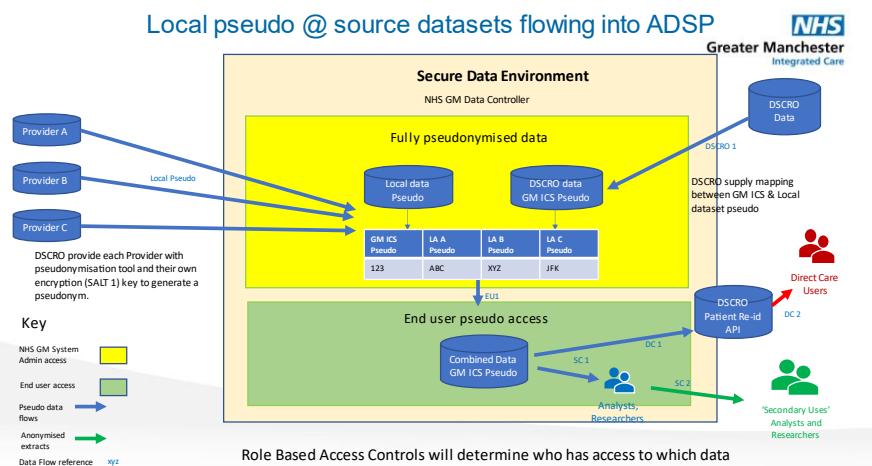
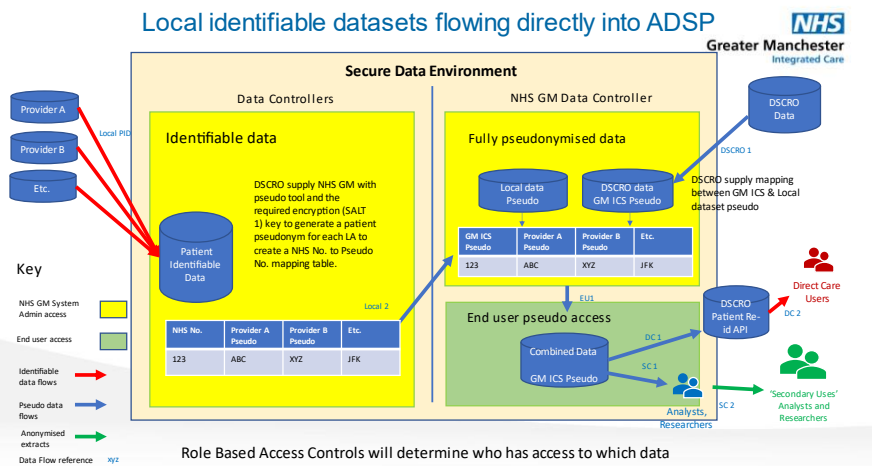
9. RBAC controls will be used to manage who has access to identifiable or pseudo data.
10. Identifiable data and 'pseudo number look ups' will be restricted to named System Administrators (NHS GM).

3. Data Flows

Data flows into the ADSP can be categorised into 4 types:

1. Identifiable datasets flowing via the GM Care Record
2. Local identifiable datasets flowing directly into the ADSP from health and care providers.
3. Local pseudo @ source datasets flowing directly into the ADSP from health and care providers.
4. National and local data flows processed and pseudonymised by AGEM DSCRO before they are passed to NHS GM ICB. These data flows are referenced as data flow 'DSCRO 1' in the diagrams below.





The table below lists the data flows

Flow Ref	Flow name/description	Going from	Going to	Method of access/transfer and control	Specify the security control(s) in place for the flow	Where will the data be stored after access/transfer?
GMCR 1	GMCR data extraction	GMCR Analytics Platform	ADSP patient identifiable	Data transfer	Access to the patient identifiable data within the GMCR Analytics and Data Science Platform is restricted to the ADSP System Administrators employed by NHS GM ICB.	GM Analytics and Data Science Platform

DPIA NHS GM Secure Data Environment – Analytics and Data Science Platform (ADSP) V1.0

					<p>Direct Azure database connection via Matillion. The data is encrypted at rest and goes via the MS Azure 'Backbone' e.g., all traffic stays within the MA Azure environment</p> <p>NHS GM ICB store the data within the patient identifiable data mart. Access to this data mart is restricted to a small number of staff working within the NHS GM Data Services Team.</p> <p>The data within the patient identifiable data mart is processed through a pseudonymisation tool to generate a NHS Number to Pseudonym mapping table.</p> <p>Access to the ADSP patient identifiable data mart is logged and monitored by the NHS GM Data Governance Lead monthly.</p>	
GMCRC 2	ADSP Pseudonymisation	ADSP Patient Identifiable view	ADSP Pseudonymised data view	Data remains within ADSP	<p>A fully pseudonymised data view is created within the ADSP which contains mapping between the GMCRC patient pseudonym and the GM ICS patient pseudonym</p> <p>Access to this data view and the pseudonym mapping table is restricted to a small number of staff working within the NHS GM ICB Data Services Team.</p> <p>Access to the data mart is logged and monitored by the NHS GM Data Governance Lead monthly</p>	GM Analytics and Data Science Platform
DSCRO 1	NHSE data flows	DSCRO Microsoft Azure Blob Store	ADSP Pseudonymised data view	Data Transfer	<p>Access to pseudonymised data within the DSCRO MS Blob Store is restricted to NHS ICB Data Service staff.</p> <p>DSCRO push data files securely to a GM Storage Account, encrypted at rest</p> <p>Access to the pseudonymised data view and the pseudonym mapping table is restricted to a small number of staff working within the NHS GM ICB Data Services Team.</p> <p>Access to the data mart is logged and monitored by the NHS GM Data Governance Lead monthly</p>	GM Analytics and Data Science Platform
EU 1	End user view	ADSP Pseudonymised data view	ADSP Pseudonymised End User view	Data remains within ADSP	<p>A fully pseudonymised data view is created within the ADSP. The view only contains the GM ICS patient pseudonym.</p> <p>Data is accessible to all staff working within the GM ICS with an approved Snowflake or Tableau user account.</p>	GM Analytics and Data Science Platform

					<p>Role Based Access Control groups are used to manage what data users can access. Patient Re-ID requests must be approved by the NHS GM Sub Licence Governance Group</p> <p>Access to the data mart is logged and monitored by the NHS GM Data Governance Lead monthly</p>	
SC 1	ADSP Secondary Use Access	ADSP Pseudonymised End User view	ADSP Pseudonymised End User view	Data remains within ADSP	<p>End users can access the fully pseudonymised data view within the ADSP via Snowflake or Tableau. The view only contains the GM ICS patient pseudonym.</p> <p>Data is accessible to all staff working within the GM ICS with an approved Snowflake or Tableau user account.</p> <p>Role Based Access Control groups are used to manage what data users can access. Patient Re-ID requests must be approved by the NHS GM Sub Licence Governance Group</p> <p>Patient level data cannot be extracted as per the NHSE Sub Licencing requirements.</p> <p>Access to the data mart is logged and monitored by the NHS GM Data Governance Lead monthly</p>	GM Analytics and Data Science Platform
SC 2	ADSP Secondary Use extract	ADSP Pseudonymised End User view	ADSP Pseudonymised End User view	Anonymised aggregate extracts	<p>End users can access the fully pseudonymised data view within the ADSP via Snowflake or Tableau. The view only contains the GM ICS patient pseudonym.</p> <p>Data is accessible to all staff working within the GM ICS with an approved Snowflake or Tableau user account.</p> <p>Role Based Access Control groups are used to manage what data users can access. Patient Re-ID requests must be approved by the NHS GM Sub Licence Governance Group</p> <p>Aggregate extracts can be taken from the ADSP.</p> <p>Access to the data mart is logged and monitored by the NHS GM Data Governance Lead monthly</p>	GM Analytics and Data Science Platform
DC 1	Direct Care request	End user request via Tableau dashboard or manual request	ADSP User Management System	Request is made within the ADSP	<p>A request is received by an end user to identify patient identifiable data to support direct care.</p> <p>All patient re-id requests must be authorised by the Data Access Committee</p> <p>The ADSP User Management system has a record of every</p>	GM Analytics and Data Science Platform

					<p>approved use case and authorised users.</p> <p>If the request to re-id satisfies the governance approvals data will be released using appropriate RBAC controls.</p> <p>Patient Identifiable data will be made available to end users using the DSCRO's patient re-id application.</p> <p>All patient re-id requests will be logged as per NHSE Sub Licencing requirements. The log will contain the reason for the request, who requested it, when the data was released, and the records made available.</p> <p>All patient re-id requests will be monitored by the NHS GM Data Governance Lead monthly.</p>	
DC 2	Direct Care access	ADSP Pseudonymised End User view	DSCRO Patient Re-ID Application	Tableau within the ADSP or an external data transfer	<p>Patient identifiable data is released to an authorised user for an approved direct care use case.</p> <p>All patient re-id requests must be authorised by the GM Secondary Uses and Research Group.</p> <p>The ADSP User Management system has a record of every approved use case and a list of authorised users.</p> <p>If the request to re-id satisfies the governance approvals data will be released using appropriate RBAC controls.</p> <p>Patient Identifiable data will be made available to end users using the DSCRO's patient re-id application.</p> <p>All patient re-id requests will be logged as per NHSE Sub Licencing requirements. The log will contain the reason for the request, who requested it, when the data was released, and the records made available.</p> <p>All patient re-id requests will be monitored by the NHS GM Data Governance Lead monthly.</p>	<p>GM Analytics and Data Science Platform</p> <p>External data transfers will need to be approved on a case-by-case basis</p>
Local PID	Local identifiable data flows direct into the ADSP	GM ICS partner source system	ADSP patient identifiable view	Data transfer	<p>Access to the patient identifiable data within the GMCR Analytics and Data Science Platform is restricted to the ADSP System Administrators employed by NHS GM ICB.</p> <ol style="list-style-type: none"> Files uploaded via a secure web portal sFTP submission Direct Blob store submission Federated Snowflake share 	GM Analytics and Data Science Platform

					<p>NHS GM ICB store the data within the patient identifiable data mart. Access to this data mart is restricted to a small number of staff working within the NHS GM Data Services Team.</p> <p>The data within the patient identifiable data mart is processed through a pseudonymisation tool to generate a NHS Number to Pseudonym mapping table.</p> <p>Access to the ADSP patient identifiable data mart is logged and monitored by the NHS GM Data Governance Lead monthly.</p>	
Local Pseudo	Local pseudonymised data flows direct into the ADSP	GM ICS partner source system	ADSP pseudonymised data view	Data transfer	<p>DSCRO issue the GM ICS partner with a pseudonymisation tool and an encryption key.</p> <p>The GM ICS partner pseudonymises and submits their encrypted data to NHS GM ICB using one of the following methods</p> <ol style="list-style-type: none"> 1 Files uploaded via a secure web portal 2 sFTP submission 3 Direct Blob store submission 4 Federated Snowflake share <p>A fully pseudonymised data view is created within the ADSP which contains mapping between the GMCR patient pseudonym and the GM ICS patient pseudonym</p> <p>Access to this data view and the pseudonym mapping table is restricted to a small number of staff working within the NHS GM ICB Data Services Team.</p> <p>Access to the data mart is logged and monitored by the NHS GM Data Governance Lead monthly</p>	GM Analytics and Data Science Platform

4. ADSP User access and management

Access to the ADSP is controlled via a single user account requiring Multi Factor Authentication (MFA). NHS GM will use the following tools to authenticate and manage user access:

Auth0: a third-party authentication and authorisation platform that supports the implementation of MFA and allows end users to have a single user account to access all applications within the ADSP.

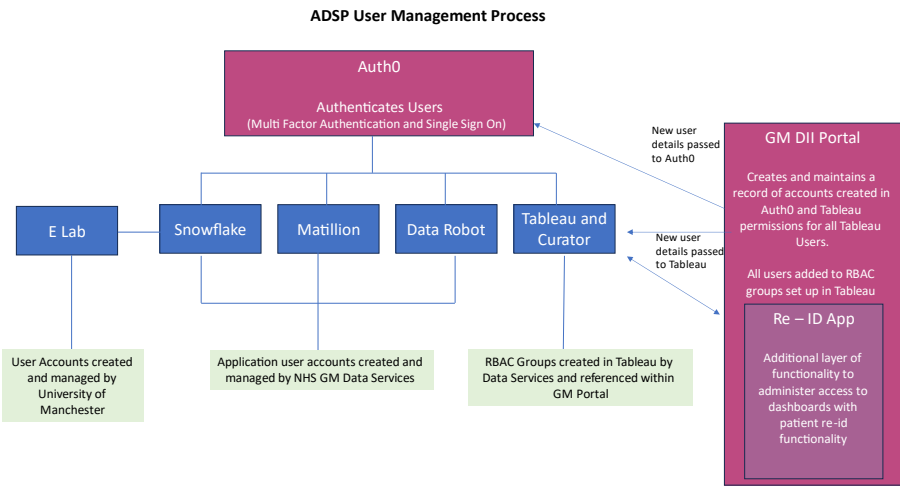
GM Data and Intelligence Portal: this is a bespoke application developed by the NHS GM Data Services and Infrastructure Team. The portal will be used by the NHS GM Data, Intelligence and Insight Team's Service Desk to create user accounts for anyone wanting to use the ADSP. All ADSP users will have a user account created within the portal.

The portal will also be used by the Service Desk to create user accounts and assign Role Based Access Control (RBAC) for Tableau. It will also be used to manage Patient Re-ID requests.

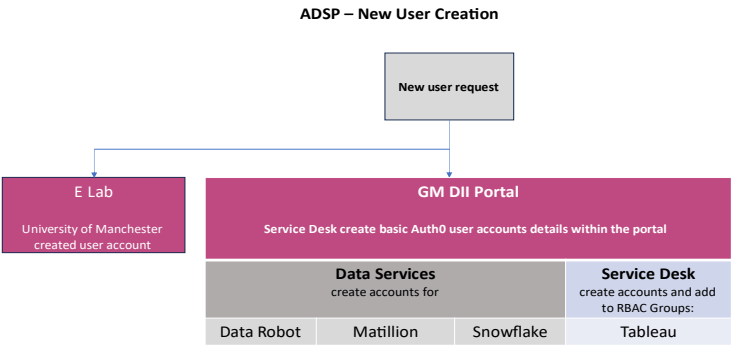
NHS GM Data Services and Infrastructure Team will assign RBAC permissions for Snowflake, Matillion and Data Robot directly within the applications.

The University of Manchester will be responsible for creating and managing access to the E Lab software.

The diagram below shows the software components and how the user management process will work.



The diagram summarises the process to add a new user:



5. Secure Data Environment - Role Based Access Control (RBAC) Groups

The following RBAC groups will be used to manage access to the SDE:

1. **System Administrator:** operating on behalf of the data controllers this is the only staff group with a single user account giving access to identifiable data and a mapping between NHS Number and the GMCR / Local patient. Small group consisting of 2 to 3 for business continuity purposes.
2. **Data Services Team:**
 - **Direct Care Group:** Small group of 8 staff will have a direct care user account giving them access to the NHS Number to GMCR / Local patient pseudonym mapping for approved direct care use cases.
 - **Secondary Use Group:** Small group of 8 staff will have a secondary uses account giving them access to the GMCR / Local to GM ICS pseudonym mapping for approved secondary use cases
3. **End user groups:**
 - **Analysts and researchers:** access to pseudonymised data, only contains the GM ICS pseudonym. No patient re-id capability
 - **Clinician or Care Professional:** access to pseudonymised data, only contains the GM ICS pseudonym with patient re-id capability for specific approved uses

Appendix D – Snowflake report



Appendix E - Secure data environment guidelines

Safe settings

The principle of 'safe setting' is about preventing inappropriate access, or misuse, of data.

The safe settings principle will be upheld by secure data environments because data security is integral to their design.

1. Secure data environments will be the default way to access NHS Health and Social Care Data for research and analysis

Secure data environments must be adopted by organisations hosting NHS health and social care data for research and analysis. These environments have features that improve data privacy and security, which will help build public trust in the use of their data.

Instances of analysing or disseminating data outside of a secure data environment will be extremely limited. Any exceptions will require significant justification, such as where explicit consent from clinical trial participants has been obtained. Further guidance about exceptions to the secure data environment standards will be provided in the coming months.

2. Secure data environments providing access to NHS health and social care data must meet defined criteria

The design, implementation and management of secure data environments must meet minimum requirements. This will include technical, behavioural, governance, and training specifications. Owners of these environments must be able to continue to demonstrate that they fulfil defined criteria in order to be categorised as an 'NHS accredited secure data environment'. All environments will be held to the same requirements and oversight.

In the coming months we will publish additional technical guidance and information governance requirements, and information about how secure data environments will be accredited. We will also communicate details about the plans, approach and timescales for this transition.

This will make sure that we can provide assurance that all NHS accredited secure data environments uphold the same privacy and security standards. It will also help to build public trust in how their data is used.

3. Secure data environments must maintain the highest level of cyber security to prevent unauthorised access to data

Secure data environments must adhere to the principle of 'security by design'. All aspects of cyber security must be integrated into the design and implementation of these environments. This includes information governance, data encryption, and data access management standards.

Security by design will make sure that secure data environments comply with the UK General Data Protection Regulation (UK GDPR) requirement of data protection by design and by default. They will uphold data protection legislation and safeguard individual rights.

4. Secure data environment owners must be transparent about how data is used within their environment

Owners of secure data environments must be open about the way data is used within their secure data environment. They must be able to detail who is accessing the data and for what purpose. This may be achieved, for example, by organisations ensuring that clear and accessible reporting is in place for their secure data environment.

The [Office for Statistics Regulation's recent report on lessons learned from the COVID-19 pandemic](#) demonstrated that public trust in the use of their data increased when they were able to see how it is used. Being transparent about how NHS health and social care data is used in secure data environments can help to build public understanding and trust.

Transparency about how data is used also increases the accountability of data controllers and data users.

Safe people

The principle of 'safe people' is about ensuring that individuals accessing data are trained and authorised, to use it appropriately.

The safe people principle will be upheld by secure data environments by making sure that users are verified before access is granted and are able to access appropriate data only. Patients and the public will also be engaged in decisions about who can access their data.

5. The secure data environment may only be accessed by appropriate, verified users

Access to NHS health and social care data within a secure data environments must be carefully controlled. Only authorised users will be granted access to data for approved purposes. Owners of secure data environments must have robust technical and governance processes in place to accurately verify the identity of users, and for managing their access to data within the environment.

This will enable a variety of users - with sufficient levels of training, qualifications, and expertise - to analyse NHS health and social care data. Allowing appropriate access to this data will facilitate data-driven planning, research, and innovation in the NHS.

6. Secure data environments must make sure that patients and the public are actively involved in the decision making processes to build trust in how their data is used

Owners of secure data environments must make sure that the public are properly informed and meaningfully involved in ongoing decisions about who can access their data and how their data is used. For example, by ensuring that relevant technical information is presented in an accessible way (that is, through publishing privacy notices and data protection impact assessments).

This will make sure that secure data environments comply with UK General Data Protection Regulation (GDPR), which requires that data controllers provide individuals with information about how their data is used.

Secure data environment owners must also be able to demonstrate that they have, or plan to, undertake active patient and public involvement activities. Patient and public involvement and engagement (PPIE) activities must follow the [NHS Research Authority's principles](#).

This guideline supports the commitments made in the Data saves lives strategy, to build and maintain public trust in the use of NHS health and social care data, through active PPIE . It will make sure that all perspectives are taken into consideration in the design and implementation of secure data environments and help build public trust in how NHS health and social care data is stored and used.

Safe data

The principle of 'safe data' is about making sure that information is protected and is treated to protect confidentiality.

The safe data principle will be upheld by secure data environments by their design and function, which prevents the dissemination of identifiable data.

7. Data made available for analysis in a secure data environment must protect patient confidentiality

Data must be treated in a secure data environment to protect confidentiality using techniques such as data minimisation and de-identification. De-identification practices mean that personal identifiers are removed from datasets to protect patient confidentiality. This includes techniques such as aggregation, anonymisation, and pseudonymisation. The level of de-identification applied to data may vary based on user roles and requirements for accessing the data.

Data minimisation practices help make sure that access to data is relevant and limited to what is necessary in relation to the purposes for which they are processed. This is in line with [Information Commissioner's Office \(ICO\) guidance](#). Applying data minimisation and de-identification practices enables approved individuals to access data for high quality analysis intended for the public good whilst also maintaining patient confidentiality.

Data protection law will continue to apply. This means there must always be a valid lawful basis for the collection and processing of personal information (including special category information) within secure data environments, as defined under data protection legislation. Where the data being accessed is confidential patient information, the requirements of the common law duty of confidentiality must also be met. More information on this can be found in the [Transformation Directorate's guidance on confidential patient information](#).

We will provide further information about the application of these practices in due course, when we publish additional guidance for secure data environments.

8. Inputs to a secure data environment must be assessed and approved

Owners of secure data environments must have robust processes in place for checking external inputs before they are approved to enter the environment. This includes data, code tools, and any other inputs.

Owners of secure data environments must have processes in place to make sure that the linking of NHS health and social care data with other datasets is carried out within the environment itself. They must also make sure that only approved and appropriately qualified individuals conduct dataset linking. This must be upheld unless there is significant justification for not doing so (in line with guideline 1).

There must also be processes in place to assure the quality of external datasets before they are imported into the secure data environment.

Linking NHS health and social care data to data from other sources has the potential to greatly enhance the quality of analysis and research findings. Secure data environments will facilitate data linkage, whilst also maintaining data protection.

Safe projects

The principle of 'safe projects' is about making sure that research projects are approved by data owners for the public good.

The safe projects principle will be upheld by secure data environments by:

- a) supporting open working practices that deliver efficiencies and improve the quality of analysis and findings
- b) making data available for a range of uses intended for the public good

9. Secure data environments must adhere to a policy of open-working and support code-sharing

Secure data environments must support open working, ensuring that code developed in these environments is reusable. Examples of how this could be achieved include:

- applying the principles of the NHS [Open Source Code policy](#)
- using the [Reproducible Analytical Pipelines \(RAP\) strategy](#)

Code developed in secure data environments must be published in the open unless there is a specific rationale for not doing so. We will engage further on these exceptions, and publish guidance in due course. This may include making it available in open repositories.

Working in the open will allow researchers to view, reuse and adapt existing code and enhance shared understanding of how the datasets in these environments are used. This will enable users to easily reproduce previous analysis, which will save time and improve the consistency and accuracy of analytical findings. This will lead to better outcomes for patients, the public, and the NHS.

10. Secure data environments must be able to support flexible and high-quality analysis for a diverse range of uses

Owners of secure data environments must engage with their intended users to make sure that they provide the necessary functionality and tools required for analysis. A range of users with different requirements and skill sets will need to access data within these environments. They will need to analyse different data to produce different outputs.

This will make sure that a variety of users will benefit from improved access to NHS health and social care data in secure data environments, which will enable data-driven planning, research, and innovation across the NHS.

11. All uses of data within secure data environments must be for the public good

The use of NHS health and social care data must be ethical, for the public good, and comply with all existing law. It must also be intended for health purposes or the promotion of health. Data access must never be provided for marketing or insurance purposes.

Owners of secure data environments must make sure there are processes in place to assess the reasons for accessing NHS health and social care data in a secure data environment. These processes must fulfil minimum national standards, which we will set out.

This will make sure that appropriate access is given to NHS health and care data, which will support the delivery of improved outcomes across the health and care system. It will also help build public confidence in why their data is accessed and how it is used.

Safe outputs

The principle of 'safe outputs' makes sure that any summarised data taken away is checked to make sure it protects privacy.

The safe outputs principle will be upheld by secure data environments by making sure that the results of analysis contain only aggregated, non-identifiable results that match the approvals of users and their projects.

12. Outputs from a secure data environment must be assessed and approved and must not identify individuals

All information must be checked before it leaves a secure data environment, including data, code, tools, and any other outputs.

There must be robust processes in place to maintain patient confidentiality and to make sure that outputs align with the intentions of individual projects.

This supports guideline 8, which states that any linking between NHS health and social care data with other datasets must be conducted within an NHS accredited secure data environment. Together these guidelines will make sure that secure data environments assist high quality analysis (for example, through data linking), whilst also maintaining data protection and patient confidentiality.

Next steps for secure data environment policy

We have now published the latest iteration of the secure data environment guidelines, expanding on the commitments made in the Data saves lives strategy. We have also published a simple explainer of secure data environment policy, which provides an outline of the policy in plain English.

The guidelines set out our ambition for secure data environment policy, which we will continue to develop in the coming months with key stakeholder groups. Below is a summary of some further planned work.

Appendix F – Direct Care Dashboards, Patient Re-Identification Tool embedded within the GM DII Portal.

A Functional Description - V1.0

Introduction

This appendix describes how NHS Greater Manchester ICB will implement the national patient re-identification tool supplied NHS England within the GM Analytics and Data Science Platform (ADSP) to support direct care.

What is Re-Id?

Patient re-identification (Re-Id) is the system's ability to take a pseudo-anonymised patient identifier and, for authorised users, return patient identifiers for the patient e.g., NHS number.

This is an important function that requires careful management and implementation to ensure that only authorised health and care professionals can re-identify patient records to support direct care. To this end, multiple controls and checks need to be implemented to ensure that re-identification is allowable and proper in various use cases.

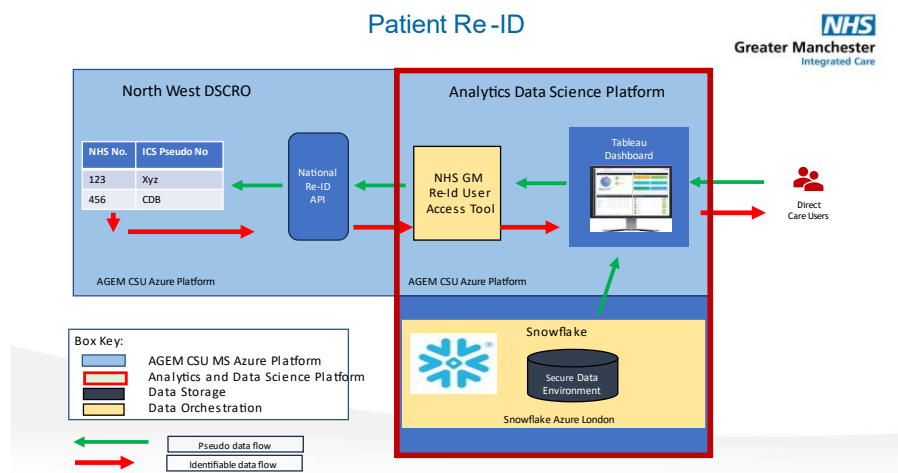
NHS GM in conjunction with Arden GEM CSU have developed a Patient Re-id tool to be used for direct care purposes. All direct care purposes e.g., Clinical Dashboards and users must be approved by the GM Data Access Committee prior to go live.

System Overview

The diagram below shows the key components of the ADSP and how the Patient Re-ID Tool works. This can be summarised as follows:

1. All data within the ADSP's Secure Data Environment is pseudonymised by default.
2. Clinical dashboards are developed within Tableau using the pseudonymised data.
3. The national Patient Re-ID application is embedded within the dashboards.
4. Authorised health and care professionals can use the Re-ID Tool functionality to request identifiers for a patient including NHS Number, Name, Date of Birth and Postcode.
5. Upon request the Re-ID User Access Tool checks the credentials of the user against the list of approved users and use cases held within the NHS GM User Access Tool.
6. If the user request satisfies the required checks a Re-ID request is made to the NHSE Data Services for Commissioners Regional Office (DSCRO) via the National Re-ID Application

Program Interface (API) to return the patient identifiers for the supplied patient pseudonym.

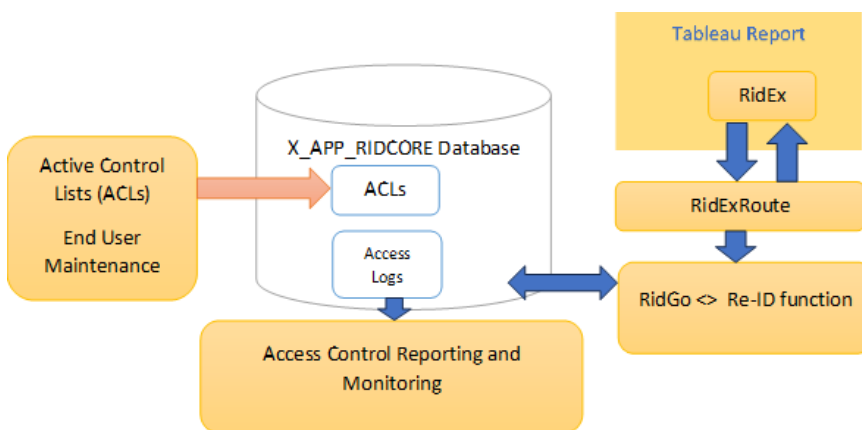


The NHS GM User Access Re-ID Tool

The diagram below and following text describe the component parts of the NHS GM User Access Re-ID Tool being developed by the NHS GM Data Intelligence and Insight Team.

An End User interface is used to record an approved list of Clinical Dashboards and End Users. The system will record a log of all Re-ID attempts and execute legitimate Re-ID requests and decline those that don't meet the required checks.

The system will provide an end user monitoring interface.



Active Control Lists (ACL) and User Maintenance

The creation and maintenance of ACLs is a non-trivial task and involves appropriate governance and a bespoke system to manage requests.

Who Will Manage ACLs?

ACLs and their associated users will need to be managed. The responsibilities of these people will be:

- Create new ACLs
- Modify existing ACLs
- Add users to ACLs
- Remove users from ACLs
- Ensure active ACLs and user configuration satisfies information governance rules.
- Be able to identify the legitimacy of ACL users and their associated sponsors.
- Analyse and act on anomalous re-id logs

The managers of the re-id system will be referred to as Access Controllers

Re-Identification Number Limit (RNL)

Each user may only be able to re-id a certain number of patients in a given time period. This is an IG limit.

At present, this limit is to be decided.

Access Control Management Tools

Managing re-id enabled users and their associate ACLs requires tools available to access controllers.

Tool Technology

No specific tool technology is defined here. We will make use of:

- Web forms
- Specialist Tableau dashboards
- SharePoint

Access Control

Access control deals with managing which users are allowed to re-id patients and which patients that they are allowed to see.

Access Control Lists (ACLs)

Each user who needs to re-id can be assigned to one or more Access Control Lists (ACLs).

Each ACL determines a cohort of patients. A user in an ACL is allowed to re-id any patients within the cohort defined by the ACL.

There are multiple types of ACL to address different types of patient cohorts.

ACL_Practice

Defined as all the patients registered at a particular GP practice. Definition is just the practice code. E.g.

ACL_Practice: P84004

ACL_ConstantCode

Patients having been assigned a specific consultant code.

ACL_NonConsultantGeog

Patients having been assigned a specialty and limited to a list of geographic areas (by postcode sectors) of patient residence.

Aimed at community groups.

E.g.

ACL_NonConsultantGeog: 710;OL9 7;OL12 9

X_APP_RIDCORE

This is a database used to store the ACLs and log each re-id attempt for security and analysis.

RidEx

The RidEx component is embedded onto a Tableau worksheet to provide re-id functionality. It provides all the user interface functions and feeds re-id requests to the processing system.

RidExRoute

RidExRoute takes each re-id request and passes it to the core re-id functions. Any high-level communication errors are identified here.

RidGo<> Routines

These functions are responsible for the real-time re-identification of authorised re-id requests coming ultimately from Tableau dashboards.

The Re-Id core performs the following functions;

1. Validates the user requesting re-id and loads their associated Access Control List(s)
2. For each id requested in the list;
 - a. Validate that this id is within the user's Access Control List(s)
 - b. Check that the user is within their pre-set RNL
 - c. If it is, return the identifiable value. If not, return the existing pseudo anonymised value.
 - d. Log the re-identification event (whether successful or not)

It is important that all the functions mentioned above are indivisible. That is, the re-id even must always be logged.

Access Control Reporting and Monitoring

Access control reporting and monitoring is important to ensure that the IG-imposed rules on re-id are being applied and that no user account is being abused.

Re-Identification API Specification

One call to the re-identification API is required per patient line to be re-identified.

The call is to a web address;

<ridCore URL>/reid/<user name>/<patientid>

E.g.

<http://127.0.0.1:5000/reid/shaun.ohara@nhs.net/8366915>

The response will be a Json frame containing the results;

```
{
  "dateOfBirth":Patient Date of Birth (time included),
  "familyName":Patient Surname,
  "givenNames":Patient Forename(s),
  "gpPracticeCode":Practice Code as Held by Re-Id Database,
  "isDeceased":Boolean - True if patient has died,
  "matchACL":Index of the Access Control List rule that allowed the Re-
id. 0 if re-id was not allowed or failed,
  "matchmessage":Textual description of re-id failure (or "OK")
  "nhsNumber":patient's NHS number
  "patientid":Original Patient Id
  "pcnCode": Practice PCN code as Held by Re-Id Database,
  "postCode":Patient postcode,
  "username":shaun.ohara@nhs.net
}
```

E.g. (Good return)

```
{
  "dateOfBirth":"1972-02-07T15:41:40.92",
  "familyName":"Bloggs",
  "givenNames":"Joe",
  "gpPracticeCode":"P112345",
  "isDeceased":false,
  "matchACL":1,
  "matchmessage":"OK",
  "nhsNumber":"9990548609",
  "patientid":"0000000001",
  "pcnCode":"123",
  "postCode":"LE3 3AF",
  "username":shaun.ohara@nhs.net
}
```

E.g. (Failed re-id)

```
{
  "dateOfBirth":"","
  "familyName":"","
```

```

    "givenNames":"","
    "gpPracticeCode":"","
    "isDeceased":false,
    "matchACL":0,
    "matchmessage":"No AGEM isMatch",
    "nhsNumber":"8366915",
    "patientid":"8366915",
    "pcnCode":"","
    "postCode":"","
    "username":shaun.ohara@nhs.net
}

```

Notice that the failed re-id returns all fields so that it can be used in any tabular report.

List of matchmessage Returns and Their Meanings

<i>Message</i>	<i>Meaning</i>
<i>OK</i>	User is authorised and re-id is returned
<i>No Match</i>	User is not authorised to re-id this patient
<i>No AGEM Return</i>	User is authorised but the re-id API call didn't return anything
<i>No AGEM isMatch</i>	User is authorised but the re-id API call indicated that there was no match