Cheshire and Merseyside Health and Care Partnership

Integrated Care Systems (ICS)

Data Protection Impact Assessment (DPIA)

Work Stream: Combined Intelligence for Population Health Action (CIPHA)

Secure Data Environment (SDE):

Sharing Data for Research with

Academia

Document Reference: ICSIGDOC-ID00010
Date agreed: October 2023
Date issued: February 2024
Next review date: February 2025

Contents

Data Protection Impact Assessment – Regional Details 5
Data Protection Impact Assessment 6
Step 1: Identify the need for a DPIA 6
Step 2: Describe the processing 8
Step 3: Consultation process 14
Step 4: Assess necessity and proportionality 17
Step 5: Identify and assess risks 25
Step 6: Identify measures to reduce risk 28
Step 7: Sign off and record outcomes 39

NHS Cheshire and Merseyside	CIPHA Combined Intelligence for Population Health Action	Health & Care Partnership for Cheshire & Merseyside

Date DPIA started:	August 2023
Date updated:	23/10/23
Next review date due by:	This DPIA will be routinely reviewed annually by the Data Asset and Access Group (DAAG) and the C&M ICS Digital and Data Information Governance Strategy Committee.
By Whom:	Suzanne Crutchley, MIAA Head of Data Protection & Information Governance Chloe Whittle, Senior Information Governance Consultant, Cheshire and Merseyside ICB Gary Leeming, Director, University of Liverpool Civic Data Cooperative
DPO approved:	Hayley Gidman, Cheshire and Merseyside ICB Data Protection Officer (DPO) 
IT Security approved:	John Llewellyn, Chief Digital and Information Officer, Cheshire and Merseyside ICB
Committee approved:	Cheshire and Merseyside ICS Digital and Data Information Governance Strategy Committee N.B. this is sign-off to the DPIA, which will then be used with the Tier Two DSA, to go out to the organisations as part of their sign-up to sharing data.
Submitted to ICO Y/N:	No

NHS Cheshire and Merseyside	CIPHA Combined Intelligence for Population Health Action	Health & Care Partnership for Cheshire & Merseyside

Information Reader Box	
Document Purpose:	Ensure consistent application of DPIA process in Work Streams
Document Name:	Data Protection Impact Assessment: Combined Intelligence for Population Health Action (CIPHA): Work Stream: **Secure Data Environment (SDE): Sharing Data for Research with Academia**
Authors:	Helen Duckworth, Director of Transformation AGEM CSU and Senior Advisor to ICB Suzanne Crutchley, MIAA Head of Data Protection & Information Governance / C&M ICS Information Governance Lead Gary Leeming, Gary Leeming, Director, University of Liverpool Civic Data Cooperative Chloe Whittle, Senior Information Governance Consultant, M&L CSU / Cheshire and Merseyside ICB
Document Origin:	Tier Two CIPHA Data Sharing Agreement
Target Audience:	All Cheshire and Merseyside Health and Care providers and commissioners as described in the: Tier Two - Data Sharing Agreement: Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE): Sharing Data for Research with Academia
Description	CIPHA Data Protection Impact Assessment for Secure Data Environment (SDE): Sharing Data for Research with Academia
Cross Reference:	SDE Principles, Tier Zero, Tier One, and the Tier Two - Data Sharing Agreement: Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE): Sharing Data for Research with Academia
Superseded Document:	CIPHA Data Protection Impact Assessment for Trusted Research Environment (TRE)
Action Required:	To note as appropriate for your organisation
Contact Details (for further information and feedback)	Chloe Whittle, Senior Information Governance Consultant, Cheshire and Merseyside ICB Information Governance Service Contact: (01782) 872648 mlcsu.ig@nhs.net
Document Status	
This is a controlled document, managed by ICB IG Strategy Group. Whilst this document may be printed, this document should not be saved onto local or other network drives.	

Data Protection Impact Assessment – Regional Details

Region	Cheshire and Merseyside	
Work Stream:	Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE): Providing Data for Research to Academia	
Version:	**February 2024**	
Reference No:	**ICSIGDOC-ID00010**	
Sharing Initiative Name:	Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE): Providing Data for Research to Academia	
Sharing Start Date:	October 2023	
Lead Organisation(s):	Cheshire and Merseyside ICB	
Work Stream Lead	**Name**	Jim Hughes
	Designation	Associate Director, Data and Digital Strategy Cheshire and Merseyside ICB
	Telephone	07788917136
	Email	Jim.hughes@imerseyside.nhs.uk
DPO Review	**Name**	Hayley Gidman
	Designation	Cheshire and Merseyside ICB Data Protection Office (DPO)
	Telephone	07809 320323
	Email	hayley.gidman@nhs.net
	Date for review	February 2025
Designated Officer Approval	**Name**	Rowan Pritchard Jones
	Designation	Executive Medical Director Cheshire & Merseyside Integrated Care System and SIRO
	Telephone	07989 570151
	Email	Rowanpj@cheshireandmerseyside.nhs.uk
	Date for review	February 2025

Data Protection Impact Assessment

Combined Intelligence for Population Health Action (CIPHA): Data Protection Impact Assessment (DPIA) - Secure Data Environment (SDE): Providing Data for Research to Academia

Step 1: Identify the need for a DPIA

Explain broadly what project aims to achieve and what type of processing it involves.
You may find it helpful to refer or link to other documents, such as a project proposal. Summarise why you identified the need for a DPIA.

National Policy Context

The Integrated Care Systems Design Framework and the "What Good Looks Like" framework 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.

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). This is a rapidly evolving policy landscape, however Cheshire and Merseyside have mature digital and data assets in place which already support the delivery of each of these national requirements.

Secure Data Environments (SDE's)

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.

The national 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 (ONS) “Five Safes Framework”, which are recognised as best practice in data protection. They are:-

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

All projects and use of CIPHA SDE data will be evaluated against these Five Safes

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 England (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 data from:

- Cheshire and Merseyside
- Greater Manchester
- Lancashire and Cumbria

The overarching purpose for data sharing is to support the Cheshire and Merseyside use of Secure Data Environments (SDEs), which are secure spaces for researchers to access pseudonymised sensitive data sets, and enables them to do data driven research, utilising the knowledge, techniques, and experience of academics to improve health. It will give researchers access to the information, which is necessary, proportionate, and relevant to their role.

Step 2: Describe the processing

Describe the nature of the processing: *how will you collect, use, store and delete data? What is the source of the data? Will you be sharing data with anyone? You might find it useful to refer to a flow diagram or other way of describing data flows. What types of processing identified as likely high risk are involved?*

The below should be read in conjunction with the **CIPHA Data Sharing Agreement: Secure Data Environment** which contains the data details; the Data Processors, Data Controllers and other information on the data sharing arrangement.

Cheshire and Merseyside have a Secure Data Environment (SDE) that houses all the patient level linked data for planning, population health and research. This environment, contracted by Cheshire and Merseyside ICB is hosted by the Data Service for Commissioners Regional Office (DSCRO) provided by Arden and GEM Commissioning Support Unit (AGEMCSU). It is intended to serve C&M ICB, Commissioner, Local Authority, provider, Voluntary, Community and Social Enterprise (VCSE) and research users.

The SDE is C&M's recognised platform for conducting all forms of analysis and research – including Population Health Management, risk stratification, planning and evaluation and research.

Technical Environment and data provisioning

The SDE is an Azure Cloud Data Management Environment. It contains several components including a cloud warehousing solution, an SQL studio query environment with various analytical and querying tools such as R, Python and also a Power BI front end to visualise data. The technical environment provisions data for projects on a project-by-project basis, ensuring data is minimised for the specific purpose. Data is provisioned in an 'Airlock' system meaning technically any analysis needs to take place within that environment and no data can leave the environment without approval and meeting certain minimisation criteria.

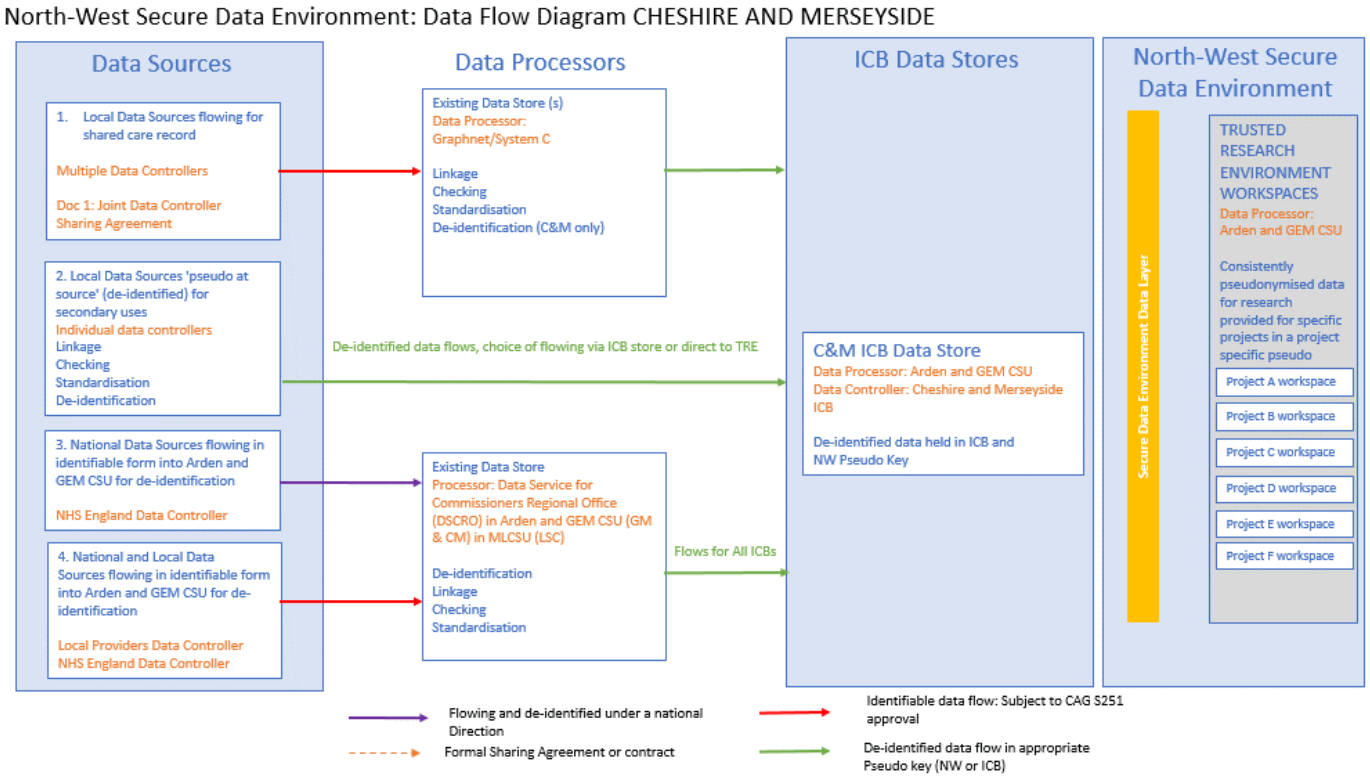
Data Flows and Data Flow diagram

Building upon recognised routes for the processing of NHSE data into AGEM CSU, the SDE combines standard national commissioning datasets with additional local data flows.

Data will flow from three core data sources:-

- National Commissioning Datasets from NHS England
- Local Organisational data flows direct from providers of services
- Data from C&M Shared Care Record provided by Graphnet as a Data Processor

Below is the data flow diagram that describes how data flows into the SDE



Pseudonymisation

Data within the Secure Data Environment will be de-identified, this means that identifiers such as name and address will be removed, Date of Birth will be converted to age and Post code will be truncated to the first 4 digits. Only pseudonymised data will be accessed by researchers within the SDE. The SDE will also contain a project specific pseudonym, so each workspace that an individual researcher accesses will have its own unique pseudonym.

Data Controllers

The Data Controllers are the C&M ICB, GP practices, NHS providers and Local Authorities in C&M.

Data Processors

The Data Processors are:

- System Supplier Graphnet Ltd using System C
- Arden and GEM Commissioning Support Unit
- Midlands and Lancashire Commissioning Support Unit.

Data Access Approval Process

Data Access approval is on a project-by-project basis. Applicants must apply to the Cheshire and Merseyside Data Access and Asset Group (DAAG) that includes nominal members from each of the Data Controllers, each of the Data Processors; a Data Protection Officer; Caldicott

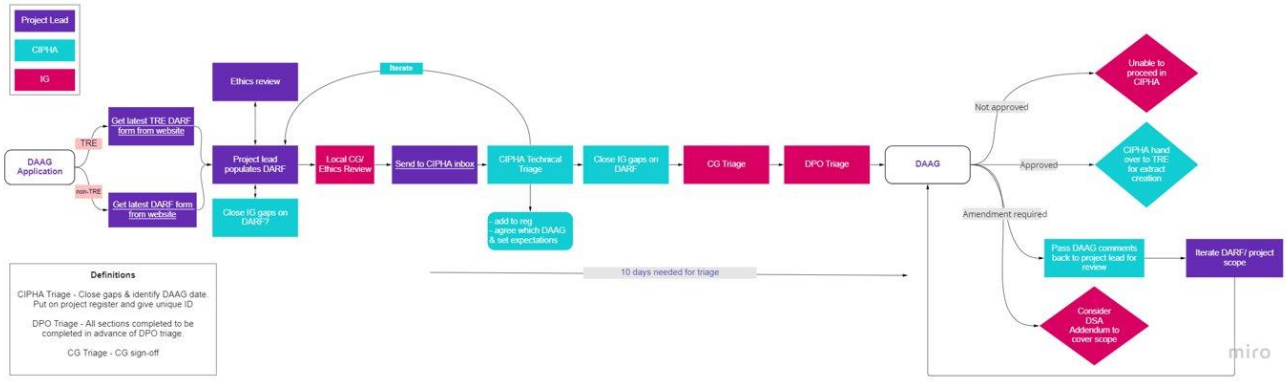
Guardian and the Public.

Applicants must complete a comprehensive Data Access Request Form (DARF) which covers questions on project scope, legal basis, datasets, data flow, Information governance, together with patient and public involvement for each project.

The process is set out below:

Diagram: Data Access Approval Process flow

N.B: TRE = Trusted Research Environment



There are two Information Governance Gateways that applicants need to achieve:-

Organisational Information Governance Gateway

The employing organisation or department within an organisation, of the individual researcher applying for the data will be required to demonstrate one of the following:

- Data Security Protection Toolkit, or
- Cyber Essentials Plus, or
- Equivalent ISO standard

Individual Information Governance Gateway

The individual will be required to undertake the Safer Research Training and sign an individual contract detailing the parameters under which the data is being used: ***CIPHA SDE Terms & Conditions of Access Agreement***, which is inserted below:

CIPHA SDE Access Agreement v1 FINAL.

Deletion and correction of information

Information can only be deleted and/or corrected by the original source organisation.

Risks/actions identified

The risks and mitigations are shown in the table in ‘Step 5’ Identify and Assess Risks (Page 25) of this document in respect of collection, storage and deletion of persistent data that is

stored within the SDE on AGEM CSU'S infrastructure.

Storage Locations

For all Data Processors (Graphnet, AGEM and ML CSU) the data is stored in the Azure Cloud:_

Microsoft Azure UK West (Backup Data Centre); Microsoft Azure UK South (Primary Data Centre)
Location Area: England & Wales
Organisation Address: Microsoft UK Headquarters, Microsoft Campus, Thames Valley Park, Reading
RG6 1WG

Describe the scope of the processing: *what is the nature of the data, and does it include special category or criminal offence data? How much data will you be collecting and using? How often? How long will you keep it? How many individuals are affected? What geographical area does it cover?*

Please see **Tier Two - Data Sharing Agreement: Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE) Sharing Data for Research with Academia Data Details**

Data shared will be for the 2.7 million individuals across Cheshire and Merseyside and also provider 'catchment' i.e. anyone who walks into a provider within Cheshire and Merseyside for care.

The data being shared includes personal sensitive health and care data. The following datasets are currently included in the Data Sharing arrangement. Others will be added as Data Controllers are added. The data flows at different intervals dependent upon dataset from daily to monthly. These are:-

National Commissioning Datasets:

For national datasets (flowing from NHS England via (DSCRO) please refer to:

https://digital.nhs.uk/services/data-services-for-commissioners/commissioning-datasets

This will always have the latest approved list of datasets.

Local Datasets are:-

- Primary care
- Acute Care shared care record
- Mental health shared care record

- Community shared care record
- North-west Ambulance Service (NWAS)
- Out of hours GP Provider flows
- Genomics (labs) data flows
- Pathology
- Radiology

Describe the context of the processing: *what is the nature of your relationship with the individuals? How much control will they have? Would they expect you to use their data in this way? Do they include children or other vulnerable groups? Are there prior concerns over this type of processing or security flaws? Is it novel in any way? What is the current state of technology in this area? Are there any current issues of public concern that you should factor in? Are you signed up to any approved code of conduct or certification scheme (once any have been approved)?*

What is the nature of your relationship with individuals?

The ICB is responsible for delivering care, improving outcomes and quality of services, and managing health and care resources for the 2.7million residents of Cheshire and Merseyside whose data will be used within the Secure Data Environment.

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

As this is national policy there are national public facing websites that describe how individual's data is used and national processes that individuals can opt out of.

Individuals can opt out of the data sharing in a variety of ways described nationally here: https://www.nhs.uk/using-the-nhs/about-the-nhs/opt-out-of-sharing-your-health-records/

In addition to these opt out routes, a local opt out will be implemented with a local contact phone line and email address for individuals to opt out on a project-by-project basis. This will be advertised as part of the communication and engagement plan which is described below.

Organisations that are party to this DSA are required to inform patients about their rights to opt-out, and are expected to also provide the public with relevant transparency and privacy notices to ensure the public is adequately informed of how health and social care organisations use their data, particularly data concerning children and vulnerable groups.

Members of the public from relevant groups are represented in the governance of the SDE specifically in the Data Asset and Access Group (DAAG) where decisions are made in respect of how data is used.

Patient Engagement

More generally Cheshire and Merseyside have a Civic Data Co-operative that regularly and consistently seeks to engage the public in how data is used. The website with multiple examples of engagement on view can be found here Civic Data Cooperative

General Communication and Engagement

In addition to the above Cheshire and Merseyside ICB have a Data and Digital Communication and Engagement plan that details the activities planned to inform the broader public of how data is used and how they can engage. More detail is within Section 3: Consultation for this DPIA.

Current State of Technology

Pseudonymisation approaches are best practice, and security and authentication are in line with best practice national policy.

Describe the purposes of the processing: *what do you want to achieve? What is the intended effect on individuals? What are the benefits of the processing – for you, and more broadly?*

The overarching purpose for data sharing is to support the Cheshire and Merseyside Secure Data Environments (SDEs), which is a secure space for planners and researchers to access health and care data. It enables them to do data driven research, utilising the knowledge, techniques, and experience of academics to improve health. These purposes are similar to the population health use cases but will be undertaken by research organisations who want to add evidence and shape policy and improve outcomes for the population of Cheshire and Merseyside.

There are five main purposes, which can be described as follows: -

Use Case 1: Epidemiology Reporting: Understanding health needs of populations, wider determinants of health and inequality for the improvement of outcomes: The data could be used to create intelligence, with the aim of understanding and improving physical and mental health outcomes, promote wellbeing and reducing health inequalities across an entire population.

Specific types of analysis that may be undertaken include:

- Health needs analysis understanding population's health outcomes and deficits;
- Demographic forecasting, disease prevalence and relationships to wider determinants of health;
- Geographic analysis and mapping, socio-demographic analysis and insight into inequalities.

Use Case 2: Predicting outcomes and population stratification of vulnerable populations: The data could be used to predict the risk of outcomes for cohorts of patients in order that services can be targeted proactively to those most vulnerable.

Use Case 3: For planning current services and understanding future service provision:

The data would be used to create intelligence on service provision to understand current service capacity and demand and forecasting future service demand to ensure enough provision is available for populations in need. This may include forecasting disease and prevalence and understanding how it impacts on service provision.

Use Case 4: For evaluation and understanding causality: The data could be used to evaluate causality between determinant of health and outcomes. Also, used to understand effectiveness of certain models of care across the health and care system.

Use Case 5: Research into novel interventions or the generation of new knowledge: The data could be used to support research into novel interventions, such as the safety of a new medication. In this case the research would be expected to generate new knowledge or to demonstrate the reproducibility of previous research.

The purposes of data sharing that are listed above are also within the document **Tier Two - Data Sharing Agreement: Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE)**

Data Details:

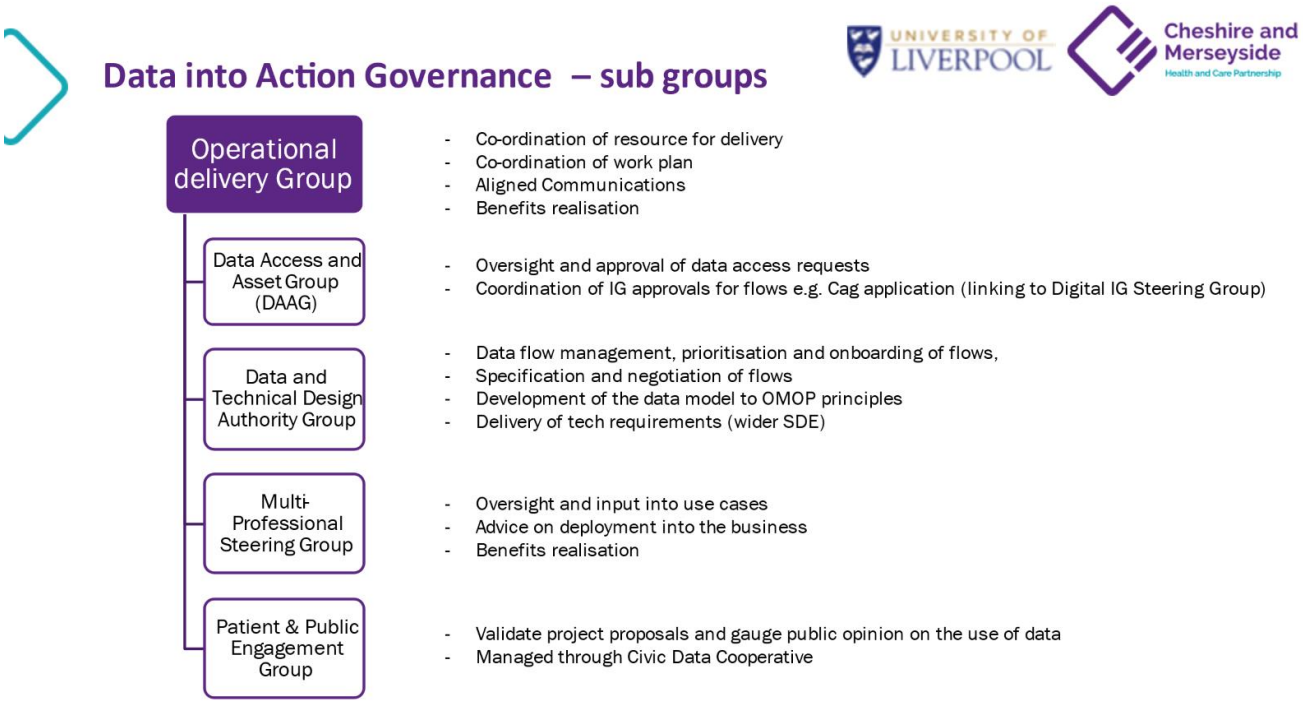
- Purpose of Data Sharing
- Details of how the Data will be shared – Data Flow

Step 3: Consultation process

Consider how to consult with relevant stakeholders: *describe when and how you will seek individuals' views – or justify why it's not appropriate to do so. Who else do you need to involve within your organisation? Do you need to ask your processors to assist? Do you plan to consult information security experts, or any other experts?*

Programme Governance

The Secure Data Environment is governed by the ICB within the Cheshire and Merseyside Data into Action Programme. There is a Data Into Action Programme Board. The sub-group structure can be found below with the following purposes. These groups are consulted on various aspects of the programme including processing and security to assure that delivery is against agreed parameters.



Data Access and Asset Group

The group that provides the gatekeeper role for information governance is the CIPHA Data Asset and Data Access Group (DAAG). This group draws its membership from information governance expertise across health and care Data Controllers, universities and providers and patient representatives. The group has a remit to ensure that requests to use the stored data for reporting maintain the integrity and purpose of the specific Data Sharing Agreement. The group will ensure the appropriateness of the Role Based Access Control (RBAC) framework in terms of individuals and groups with access to the shared record. The groups functions from the terms of reference are below:-

- Oversight of the Data Access Request Process, approving data access requests from organisations
- Ensuring all Information Governance requirements are met including GDPR and the Common Law Duty of confidentiality, Caldicott principals, data minimisation and public benefit.
- Ensuring programmes applying have undertaken appropriate Patient and Public Involvement and Engagement in their design
- Ensuring individuals and organisations applying have met the required conditions for data access
- Ensuring technical specifications of data and technology are accurate and data is minimised to project specific requirements
- Oversight of the development of the process, ensuring the process is developed in line with any changes to national policy or data sharing arrangements and escalating to the Information Governance Sub-committee where changes need to happen
- Ensuring Applications have adequate scientific critique of research proposals

Patient Engagement

More generally Cheshire and Merseyside have a Civic Data Co-operative that regularly and consistently seeks to engage the public in how data is used. The website with multiple examples of engagement on view is found here Civic Data Cooperative

General Communication and Engagement

In addition to the above Cheshire and Merseyside ICB have a Data and Digital Communication and Engagement Plan that details the activities planned to inform the broader public of how data is used and how they can engage.

- Homepage: About the SDE and why access to data is vital for research.
- FAQs.
- Research projects: What current projects the SDE is linked to and aims of the projects.
- What we do with your data and National and Local Data Opt Out.
- Organisations involved in the SDE (with links to some of the relevant websites).
- What is happening in C&M (with links to the local ICB websites):
- Privacy Notice.
- How to get involved.
- News and Events.
- Extranet for staff (if required).

The following are also planned:

Public facing engagement events: To raise the awareness of the SDE, data sharing, the importance of access to data and the right to opt out. Two half-day face to face events with key speakers from the programme.

Staff engagement events: To ensure that staff at partner organisations are aware of the programme and can actively promote the benefits through their day-to-day contact with patients/service users/public a staff awareness raising campaign will be developed. Two half-day face to face events with key speakers from the programme. (Could be virtual to save costs and allow flexibility)

Centralised suite of communications and engagement collateral: To ensure there is a consistent message across the Cheshire and Merseyside ICS a communications toolkit for the SDE will be developed by the communications and engagement campaign team. To include, but not limited to:

- Website and extranet content.
- Public facing materials – poster/leaflet/social media messaging/FAQs/patient user stories.
- Staff engagement toolkit - a presentation for team managers to utilise with their teams, a guide of what and how to share the SDE with patients/service users, messaging for internal social media channels, FAQs, training materials for staff, branding and templates.
- Animations/videos

- Programme updates.
- End of campaign newsletter highlighting progress and what next.

Data Processors

Data Processors named in this agreement are governed by Data Processing Agreements which detail the processing instructions including responsibilities around security and Information Governance. Please see next section for more detail on this.

Step 4: Assess necessity and proportionality

Describe compliance and proportionality measures, in particular: *what is your lawful basis for processing? Does the processing actually achieve your purpose? Is there another way to achieve the same outcome? How will you prevent function creep? How will you ensure data quality and data minimisation? What information will you give individuals? How will you help to support their rights? What measures do you take to ensure processors comply? How do you safeguard any international transfers?*

Legal Basis under the General Data Protection Regulation (UK GDPR)

Below is the lawful basis for processing under GDPR.

Processing Personal Data - Article 6
6(1)(e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

Processing Sensitive Personal Data – Article 9
9(2)(h) processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of domestic law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3;

9(2)(j) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) (as supplemented by section 19 of the 2018 Act) based on domestic law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

Common Law Duty of Confidentiality

For Research the Common Law Duty of Confidentiality requires that there should be no use or disclosure of any confidential patient information for any purpose other than the direct clinical care of the patient to whom it relates, unless:

- The patient explicitly consents to the use or disclosure;
- The disclosure is required by law;
- The disclosure is permitted under a statutory process that sets aside the duty of confidentiality.

For local flows coming from the Graphnet/System C supplied shared care record, data is de-identified and pseudonymised under a CAG approved Section 251 for using data for research.

National Data is pseudonymised within Data Service for Commissioners Regional Office (DSCRO) which is subject to a CAG application for S251 to be used for the purpose of research.

Data is appropriately de-identified (pseudonymised) and therefore at the point of access by planners and researchers is not owed a duty of confidentiality.

Alternative to achieve the same outcome

The alternative to this outcome is to seek consent from 2.6 million individuals which would be resource intensive and is not viable given that a large proportion may not respond. Pseudonymisation and de-identification is the suggested approach.

Data Sharing Agreements and Sub-licencing agreements

Scope is defined and scope creep is minimised by the use of Data Sharing Agreements and sub-licencing agreements that clearly define the parameters under which the data is being shared.

National NHS England Commissioning Datasets are governed by a Data Sharing Agreement between Cheshire and Merseyside Integrated Care Board and NHS England.

Local data flows via Graphnet shared care record are covered by a local CIPHA Data Sharing Agreement between each service provider Data Controller within Cheshire and Merseyside ICB and Cheshire and Merseyside ICB. This is a Joint Data Controller arrangement.

Data is sub-licenced from these two agreements by Cheshire and Merseyside ICB to the organisation employing the researcher. The sub-licence describes the conditions under which the data is accessed.

Data Processing Agreements

Measures are taken to ensure Data Processors comply by setting out data processing requirements within Data Processing Agreements.

Graphnet/System C Data Processing Agreement

STHK CCN Data Processing Agreeme

Arden and GEM CSU Data Processing Agreement

Data Minimisation in SDE

The data used for research is only accessible in the following way:

- All data is pseudonymised such that no user of the SDE could re-identify any citizen
- No record level data ever leaves the SDE
- Only anonymised data, such as diagrams, charts or aggregated tables, is extracted from the SDE for use in research papers or outputs

Any deviations in project scope that result from:

- A change in data processing responsibilities
- A change in storage, transmission, and/or persistence of data
- A change from read-only to write-back
- A change in data details from the Tier Two documentation
- A change in system architecture

will prompt a review of this DPIA in advance of the set review date, to ensure that data processing remains lawful.

Publishing of Research

The researchers will have access to de-identified, pseudonymised individual level data within the Secure Data Environment (SDE). The risk of re-identification is further reduced by only permitting access to the data specifically required to support the research question being investigated. Individual level data will not be removed from the SDE.

All researchers who access the data within the SDE will need to demonstrate that they have completed safe researcher training and are an accredited researcher, such as the course provided by the Office for National Statistics, before being permitted to access the data. See:

Become an accredited researcher - Office for National Statistics (ons.gov.uk)

De-identification and anonymisation of the data will be in line with the NHS England and ICO standards. Statistical disclosure checks are required to ensure that any data, such as charts, diagrams or tables, are safe to export. This process will be audited and will include anonymisation techniques as suggested by the UK Data Service [Anonymising quantitative data — UK Data Service]. See:

https://ukdataservice.ac.uk/learning-hub/research-data-management/anonymisation/anonymising-qualitative-data/

This includes a list of primary anonymisation techniques as follows:

- Remove direct identifiers;
- Aggregate or reduce the precision of a variable;
- Generalise the meaning;
- Restrict the upper or lower ranges;
- Anonymise relational data;
- Anonymise geo-referenced data

Once this aggregated data has egressed from the SDE it will be used for publication in academic papers, and to support other uses, such as determining health and care policy decisions, or the efficacy of new interventions.

Information given to Individuals

Cheshire and Merseyside ICB have a Data and Digital Communication and Engagement Plan that details the activities planned to inform the broader public and Data Controllers of how data is used and how they can engage this includes updates to the public facing website. See Section 3: Consultation in this document for the more detailed plan.

Re-use of Data Sets

A key concept in modern data science is known as FAIR - that data should be: Findable; Accessible; Interoperable; Re-usable. This is available at:

FAIR Principles | Library Research Support (open.ac.uk)

This is important for both the validation of research as well as improving efficiency / reducing waste through time spent on cleaning and validation of datasets. These principles do not mean that data should be open for all to use, but that there should be clear processes for archiving, re- use and publication of metadata that describes the dataset.

Datasets used within the SDE will be archived and metadata published within the CIPHA metadata catalogue. Further use and access of this data will be subject to further approval based on the same processes as for new data applications.

Data Processors Responsibilities to the Public

In the event that personal information which has been shared under the DPIA is compromised or possibly compromised, the agency making the discovery will without delay:

- Inform the organisation providing the details within 24 hours
- Take steps to investigate the cause
- Report and investigate as an incident
- If appropriate, take disciplinary action against the person(s) responsible
- Take appropriate steps to avoid a repetition.

On being notified that an individual's personal information has or may have been compromised, the original provider will assess the potential implications for the individuals whose information has been compromised, and will:

- Notify the individual concerned
- Advise the individual of their rights
- Provide the individual with appropriate support
- Undertake a generalised risk assessment and consider notifying the Information Commissioner's Office (ICO) in line with expected procedure

Data Protection Assessment

<table>
<tr><td>Data Protection Principal</td><td>Data Protection Principal</td><td>Comments</td></tr>
<tr><td>Lawfulness, fairness and transparency</td><td>Lawful Basis</td><td>Legal Basis under the General Data Protection Regulation (UK GDPR)

Below explains how this SDE work is compliant with UK GDPR:

Processing Personal Data - Article 6
6(1)(e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

Processing Sensitive Personal Data – Article 9
9(2)(h) processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of domestic law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3;

9(2)(j) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) (as supplemented by section 19 of the 2018 Act) based on domestic law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

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)
Official authority:

<table><tr><td>GP Practices</td><td>NHS England’s powers to commission health services under the NHS Act 2006. Also, Article 6 (1) c for GPs when subject to statutory regulation</td></tr></table></td></tr>
</table>

		NHS Trusts	National Health Service and Community Care Act 1990
		NHS Foundation Trusts	Health and Social Care (Community Health and Standards) Act 2003
		Local Authorities	Local Government Act 1974 Localism Act 2011 Children Act 1989 Children Act 2004 Care Act 2014
	Fairness	Individuals can exercise the following rights with respect to their data, where applicable, by contacting the source organisation of their data: • Right of access • Right to rectification • Right to erasure • Right to restrict processing • Right to data portability • Right to object • Rights related to automated decision making • Rights related to including profiling	
		For Research the Common Law Duty of Confidentiality requires that there should be no use or disclosure of any confidential patient information for any purpose other than the direct clinical care of the patient to whom it relates, unless: •The patient explicitly consents to the use or disclosure; •The disclosure is required by law; •The disclosure is permitted under a statutory process that sets aside the duty of confidentiality. Appropriately pseudonymised, de-identified or aggregated data is not owed a duty of confidentiality. Under this Data Sharing Agreement the Common Law Duty of Confidentiality does not apply, as the data is pseudonymised, and presented as aggregate data.	
	Transparency	The responsibility for transparency lies firmly with the Data Controllers who are the partner organisations within the CIPHA Work Stream.	
Purpose limitation		Combined Intelligence for Population Health Action (CIPHA): Secure Data Environment (SDE)	
Research		The purposes of data for local Intelligence Services are those described in the UK GDPR Article 9(2) (j) – research.	

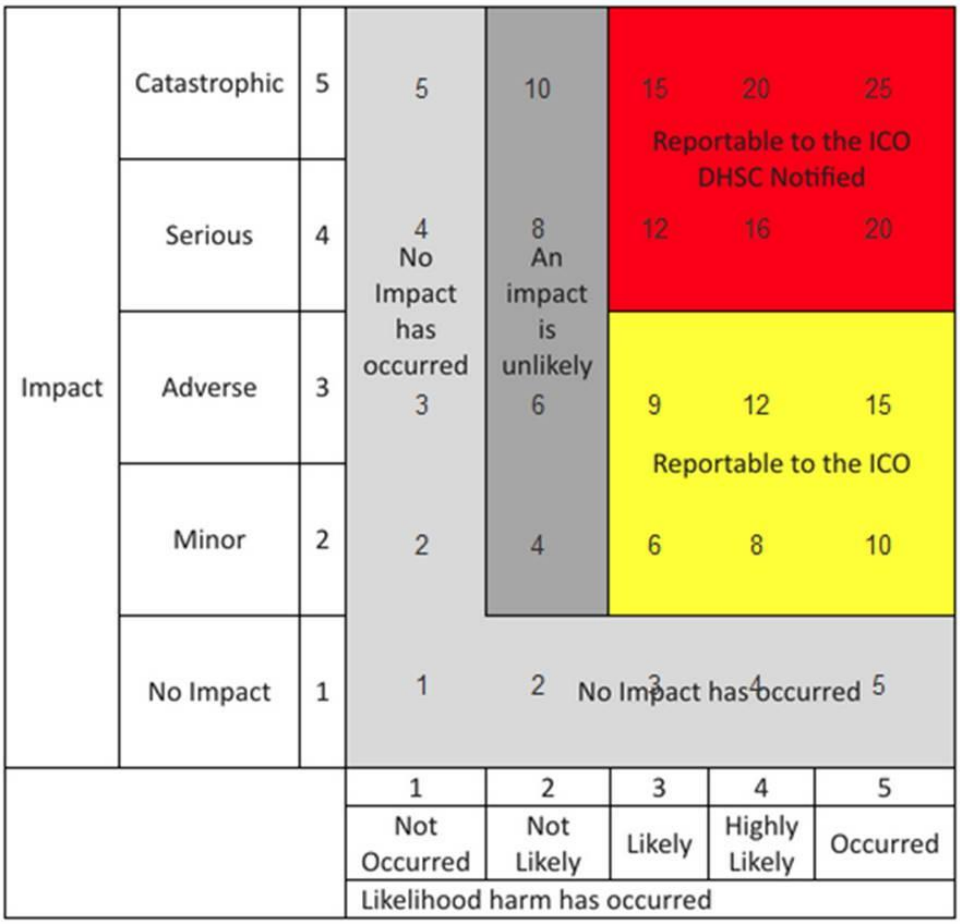
Data Minimisation	**Sensitive Codes / legally restricted** Legally restricted codes were previously referred to as sensitive codes. Sensitive data excluded from retrieval follows the recommendations made by The Royal College of General Practitioners (RCGP) Ethics Committee and the Joint GP IT Committee: • Gender reassignment. • Assisted conception and in vitro fertilisation (IVF) • Sexually transmitted diseases (STD) • Termination of pregnancy For data from local authorities some special category/sensitive data is included, and the inclusion is covered by the legal basis for sharing. All free text data fields are omitted from data collection.
Accuracy	Data Quality Improvement Plans exist for all data assets within the Secure Data Environment
Storage limitation	The data will be stored in line with the NHS Records Management Code of Practice 2021 A guide to the management of health and care records
Integrity and confidentiality	Access levels to information available through AGEM CSU will be based upon the role held by the provider of health and care. Information will be shared which is necessary, relevant and proportionate to the role the individual fulfils.

Role Based Access Control (RBAC)	**Role Based Access Controls** Role Based Access Controls (RBAC) will be applied to a Secure Data Environment (SDE). The SDE will sit on AGEM CSU technical infrastructure (named as a Data Processor to this agreement). The RBAC are as follows: - **Patient Identifiable Data:** Access to patient identifiable data **will not** be granted as part of this Data Sharing Agreement. **Pseudonymised Data:** Research organisations and individuals from those organisations, who are approved via an Information Governance Gateway, described in the '**Governance'** section, will be granted access to a sub- set of pseudonymised data relating to their research specification and in line with the purpose of data sharing. All research projects undertaken with the data will be listed in the Data Access and Asset Matrix explaining the specific purpose, requestor, output, legal basis, and timescale. **Anonymised-aggregate Data:** Organisations and individuals from those organisations, who are approved via an Information Governance Gateway, described in the '**Governance'** section, will be granted access to anonymised-aggregate data via software such as OpenSAFELY. This will only be for data relating to their research specification and in line with the purposes of data sharing. All research projects undertaken with the data will be listed in the Data Access and Asset Matrix explaining the specific purpose, requestor, output, legal basis, and timescale.

Step 5: Identify and assess risks

CIPHA Risk Log

The risk score uses the following matrix:



No.	Effect	Description
1	No adverse effect	There is absolute certainty that no adverse effect can arise from the breach
2	Potentially some minor adverse effect or any incident involving vulnerable groups even if no adverse effect occurred	A minor adverse effect must be selected where there is no absolute certainty. A minor adverse effect may be the cancellation of a procedure but does not involve any additional suffering. It may also include possible inconvenience to those who need the data to do their job.
3	Potentially some adverse effect	An adverse effect may be release of confidential information into the public domain leading to embarrassment or it prevents someone from doing their job such as a cancelled procedure that has the potential of prolonging suffering but does not lead to a decline in health.
4	Potentially Pain and suffering/ financial loss	There has been reported suffering and decline in health arising from the breach or there has been some financial detriment occurred. Loss of bank details leading to loss of funds. There is a loss of employment.
5	Death/ catastrophic event.	A person dies or suffers a catastrophic occurrence

Risk Number	Describe source of risk and nature of potential impact on individuals.	Likelihood	Impact	Overall Risk Score
1.	That data is not adequate to link records appropriately or sufficiently well coded for accuracy the consequence being that the findings drawn from the analytics are thus diluted.	Not likely	Serious	**8**
2.	Failure to keep clients informed over how their data will be used could lead to a breach of UK GDPR Article 13 and 14 of the GDPR.	Not Likely	Serious	**8**
3.	Failure to have processes in place to facilitate the following data protection rights requests could result in a breach Article 15, Article 16, Article 18 and Article 21 • Right of Access • Right to Rectification • Right to Restrict Processing • Right to Object Under this Data Sharing Agreement, the data being accessed will be in aggregated or in a consistently pseudonymised form.	Not Likely	Serious	**8**
4.	Failure to ensure that the supplier is compliant with Government and National Cyber Security Standards for cloud based computing could lead to a breach of our security obligations under Article 32 of the UK GDPR.	Not Likely	Serious	**8**
5.	Failure to define the process in which direct care providers outside of an LA area can access the records of patients outside of their area could result in data being accessed inappropriately leading to a breach of each partner's security obligations under Article 32 of the UK GDPR.	Not Likely	Catastrophic	**10**
6.	Failure to have security processes in place to stop partners, with access to patient identifiable data, from accessing the portal from their own personal devices, this could result in a breach of each partner's security obligations under Article 32 of the UK GDPR.	Not Likely	Catastrophic	**10**
7.	Failure to have a process in place to audit access to patient identifiable data processes could result in a breach of our security obligations under Article 32.	Not Likely	Serious	**8**

8.	Failure to ensure adequate controls are in place to ensure that de-identified data can't be re-identified could result in disclosure of personal information leading to a data breach and could lead to a breach of our security obligations in relation to anonymisation / pseudonymisation processes under Article 32.	Not likely	Catastrophic	**10**
9.	Failure to have a process in place to verify, audit and test the merging of data from multiple data sources to ensure that data is matched correctly to ensure that a data breach does not occur.	Not likely	Catastrophic	**10**
10.	Failure to provide / develop a process / technical solution to facilitate clients opting out of their data being shared could lead to a breach of the Common Law Duty of Confidentiality, Data Protection Act and Human Rights Act.	Not Likely	Catastrophic	**10**
11.	Failure to ensure that a process is in place to remove a client's data when the partner has closed the record on their systems could result in data being retained inappropriately.	Not Likely	Catastrophic	**10**
12.	Failure to ensure that the appropriate international transfer safeguards are in place should the note data be stored on servers outside of the UK could result in a breach of Article 44-56.	Not likely	Catastrophic	**10**
13.	Failure to define the retention of closed records data on the system could result be held on the portal inappropriately.	Not Likely	Catastrophic	**10**

Step 6: Identify measures to reduce risk

All risk mitigations to be checked/amended/deleted/added in as appropriate

Risk Number	Risk Summary	Options to reduce or eliminate risk	Residual Risk: Low Medium, High	Effect on Risk: Eliminated, Reduced, Accepted	Measure Accepted: Yes/No
1.	That data is not adequate to link records appropriately or sufficiently well coded for accuracy the consequence being that the findings drawn from the analytics are thus diluted.	To use operational flows where possible which reflect actual activity and both in the testing and regular feedback that data quality is given due attention and resource to resolve issues that arise. Routine data quality reports will be available e.g. “orphan” activity records by provider that will be applied to business-as-usual governance.	**Low**	Reduced	Yes
2.	Failure to keep clients informed over how their data will be used could lead to a breach of UK GDPR Article 13 and 14 of the GDPR.	and the legal basis that is being used to share data.	**Low**	Reduced	Yes

3.	Failure to have processes in place to facilitate the following data protection rights requests could result in a breach Article 15, Article 16, Article 18 and Article 21 • Right of Access • Right to Rectification • Right to Restrict Processing • Right to Object Under this Data Sharing Agreement, the data being accessed will be in aggregated or in a consistently pseudonymised form.	Each Data Controller is accountable under UK GDPR and will have their own measures in place to meet the eight Rights of Data Subjects. If a Data Subject of any partner organisation wishes to exercise or challenge one of their Rights, they would do that with their provider organisation(s) through the partner organisation's internal processes. Each Data Controller will remain responsible and accountable under UK GDPR for their clients. The host Trust of the platform – Mersey and West Lancs Hospitals NHS Trust – have in place their data processing and cyber policies and procedures to maintain the rights of the data subjects.	**Low**	Reduced	Yes
4.	Failure to ensure that the supplier is compliant with Government and National Cyber Security Standards for cloud based computing could lead to a breach of our security obligations under Article 32 of the UK GDPR	Data will be stored on 'Azure cloud', which is compliant with Information Governance standards and is safe and secure. Azure is assessed to ISO 27001, ISO 27017, ISO 27018, and many other internationally recognized standards. The scope and proof of certification and assessment reports are published on the Azure Trust CenSDE section for ISO certification here: https://www.microsoft.com/en-	**Low**	Reduced	Yes

		us/trustcenter/compliance/iso-iec27001. The ISO 27001 assessment was performed by the BSI. SystemC and Graphnet Health Ltd comply with the 13 Infrastructure as a Service (IaaS) principles and are accredited as such e.g. Cyber essentials. Details are available on request contained within the "CareCentric population health cloud assurance" document.			
5.	Failure to define the process in which direct care providers outside of an LA area can access the records of patients outside of their area could result in data being accessed inappropriately leading to a Data Protection Act Section 170 offence	The following processes are in place • The supplier defines rigorous role-based access (RBAC) protocols to ensure access to data is limited to those authorised and maintains a register of RBAC • The supplier maintains an audit trail of access to data sources • The Work Stream controls access to data assets through a 'Data Asset and Access Group' to ensure only legitimate access is granted to individual projects (use-cases). This is linked to the RBAC process	**Low**	Reduced	Yes
6.	Failure to have security processes in place to stop partners, with access to patient identifiable data, from accessing the portal from their own personal devices, this could result in	The following mitigating processes are in place • Personal identifiable data can only be made available (re-identified) using the existing and approved pseudonymised at source' mechanism. Be this direct or via virtual private network (VPN)	**Low**	Reduced	Yes

	a breach of each partner's security obligations under Article 32 of the UK GDPR	will be subject to the acceptable usage policy of the organisation that the person making access works for. Each individual will be subject to the policies and procedures outlined by their employer			

7.	Failure to have a process in place to audit access to patient identifiable data processes could result in a breach of our security obligations under Article 32.	The following mitigations are in place; • The Azure SQL environment logs all SQL queries which take place against the data marts to provide an audit trial of what identifiable data has been accessed and by whom • Requests for re-identification of cohorts through the Web Client application are recorded separately and will be provided on a regular basis to the CIPHA board • Access to the data will be subject to approval from the Data Controllers. The existing change control process would approve access and grant permissions • All activity reports are available as outlined above and would be provided to assist audit. Audit process and timeframes will be specific to each organisation The Work Stream controls access to data assets through a 'Data Asset and Access Group' (DAAG) to ensure only legitimate access is granted to individual projects (use- cases).	**Low**	Reduced	Yes

8.	Failure to ensure adequate controls are in place to ensure that de-identified data can't be re-identified could result in disclosure of personal information leading to a data breach and could lead to a breach of our security obligations in relation to anonymisation / pseudonymisation processes under Article 32	Direct Care data marts hold the full PID along with field level configuration for both anonymisation and legally restricted clinical coding reference data. Stored procedures query tables using filed level configuration to anonymise data at the point of extract. SSIS package cross references data with legally restricted clinical coding to further remove restricted data. Fully anonymised data is written to the research data mart in the same format as the direct care source. Key masking uses a customer specific SALT value + SHA2_256 hashing. **Security** • Separate cloud security helpdesk with one request per user • IP addresses must be whitelisted for access to data marts • Azure AD named user access must be used • Data access can be controlled by mirroring CareCentric RBAC configuration • Full SQL row level security • Unique RBAC groups can be implemented within analytics solution if required **Anonymisation** • Source is the Direct Care mart holding all data • Data is copied to the Anonymised mart • Legally restricted clinical codes stripped out in flight	**Low**	Reduced	Yes

		• Field level configuration for anonymisation
		  • No change
		  • Blank
		  • Truncate
		  • Mask Dates
		• Key fields undergo one way encryption, maintaining referential integrity
		
		Pseudonymisation
		
		• Source is the Direct Care mart holding all data
		• Data is copied to the Pseudonymised mart
		• Opted Out patients and legally restricted clinical codes stripped out in flight
		• Field level configuration for Pseudonymisation
		  ○ No change
		  ○ Blank
		  ○ Truncate
		  ○ Mask Dates
		• Tokenised IDs Can be re identified
		  ○ National DE ID / RE ID or encrypted local values
		  ○ Secured data table which stores mapping
		  ○ User interface to reidentify
		• Key fields undergo two-way encryption, maintaining referential integrity
		
		A white box penetration test has been completed with a Black box full test scheduled for 2020.			

9.	Failure to have a process in place to verify, audit and test the merging of data from multiple data sources to ensure that data is matched correctly to ensure that a data breach does not occur	Graphnet merges data into its longitudinal patient record based on the patient NHS Number, name and date of birth.	**Low**	Reduced	Yes
		Where the NHS number is a verified number, we would match on this. If this is not the case, we use the three items described above. Reports are available that outline the match success and Graphnet have performed audits for clients to ensure data integrity. The tools available to client are designed to support the ongoing data quality process which is the responsibility of each Data Controller.			

10.	Failure to provide / develop a process / technical solution to facilitate clients opting out of their data being shared could lead to a breach of the Common Law Duty of Confidentiality, Data Protection Act and Human Rights Act	Type 1 opt-outs (those who do not want their information shared outside of General Practice for purposes other than direct care) will be upheld. This means that data for people who have objected to sharing their data will not flow from the GP record into the Graphnet solution. Once the national solution for opt out is live with NHSD, these patients will automatically be removed from the datamart. This removal includes all data sources. The ability to opt out for direct patient care would only be instigated subject to a successful application by the data subject under Article 21 of UK GDPR.	**Low**	Eliminated	Yes

11.	Failure to ensure that a process is in place to remove a client's data when the partner has closed the record on their systems could result in data being retained inappropriately	The **NHS Records Management Code of Practice 2021** **A guide to the management of health and care record** sets out what people working with or in NHS organisations in England need to do to manage records correctly. It's based on current legal requirements and professional best practice. All organisations must ensure compliance with Article 5(1)(e) of the GDPR. Each organisation will abide by their own document retention/ destruction policy Each organisation will have its own records management policy and define both the duration of retentions and removal policy. The Data Processor will hold data in line with the contract terms. All data will be returned and purged at contract end, or as set out in the contractual terms.	**Low**	Reduced	Yes
12.	Failure to ensure that the appropriate international transfer safeguards are in place should the note data be stored on servers outside of the UK could result in a breach of Article 44-56	The supplier, Graphnet Health, are a UK based company. All data is stored in the UK and there is no server storage outside of the UK. All information can be found in the CareCentric population health cloud assurance document.	**Low**	Eliminated	Yes

13.	Failure to define the retention of closed records data on the system could result be held on the portal inappropriately	The **NHS Records Management Code of Practice 2021** **A guide to the management of health and care record** sets out what people working with or in NHS organisations in England need to do to manage records correctly. It's based on current legal requirements and professional best practice. Each organisation that contributes to the solution will have a record retention policy. The elements of the record, when combined, creates a holistic view of a care recipient's journey. As a result, this new record would be retained for the duration of the longest term for which the record is retained within the social care community.	**Low**	Reduced	Yes

Step 7: Sign off and record outcomes

Item	Name/date	Notes
Measures approved by:	Helen Duckworth 23/10/23	Final approval by: C&M ICS Digital and Data Information Governance Strategy Committee
Residual risks approved by:	Jim Hughes 23/10/23	Final approval by: C&M ICS Digital and Data Information Governance Strategy Committee
DPO advice provided:	Suzanne Crutchley 23/10/23	DPO should advise on compliance, step 6 measures and whether processing can proceed
DPO Comments: This work for **the Combined Intelligence for Population Health Action (CIPHA) Secure Data Environment (SDE)** meets the requirements for UK GDPR, and so the data processing can proceed.		
DPO advice accepted or overruled by:	Jim Hughes 23/10/23	If overruled, you must explain your reasons
Comments: Approved		
Consultation responses reviewed by:	Suzanne Crutchley and Helen Duckworth 23/10/23	If your decision departs from individuals' views, you must explain your reasons
Comments: September 2023: Consultation for feedback/comments/amendments to: ✓ C&M IG SIGN Members ✓ C&M ICS DD IGSC Members Documents: • Data Sharing Agreement Tier Two: Secure Data Environment (SDE): Research • DPIA: Secure Data Environment (SDE): Research • Cheshire and Merseyside Secure Data Environment: FAQs		
This DPIA will be kept under review by:	The C&M ICS CIPHA: DAAG, and approval by the C&M ICS DDIGSC Annually – on or before February 2025	The DPO should also review ongoing compliance with DPIA

Please return to: mlcsu.ig@nhs.net