Appendix to the Report: Enabling data flows in Greater Manchester Connected Health City

Contents

CHI TRE Project Application Form

Project No. (provisional, to be filled in by TRE Operations team):	
Date form received (to be filled in by TRE Operations team):	

This form should be completed by researchers who wish to host a project within the University of Manchester (UoM) Centre for Health Informatics (CHI) Trustworthy Research Environment (TRE). Its purpose is to provide the TRE Project Board with the necessary information to review and approve proposed projects.

Please email your completed form and supporting documentation, e.g. ethical approval letter or Data Sharing Agreements, to: tre-support@manchester.ac.uk

Part A. Project Details

1. **Project Title** (as chosen by the project team/funder) – *75 characters maximum*

2. **Project Application name if different from above** (as used by the Data Controller, e.g. NHS Digital) - *75 characters maximum*

3. **Project Summary** (include the aims of the project, name of funder, and what data flows are involved) – *300 characters maximum*

4. **Does your project have any funding to cover TRE resources?** (e.g. staff time for technical or information governance support) **Yes / No**
   If yes, please provide details below:

5. **Is your project part of a larger programme: Yes / No**
   If yes, please give details of the programme, particularly if it comprises other projects that might need to use the TRE:

6. **Data Requirements**

6.1 Will your project analyse data as part of a research or service evaluation project? Yes / No

If yes, such projects typically get access to one dedicated virtual machine. Please indicate your requirements for this machine

TRE Resource type	Details
Virtual workstation (Linux or Windows?)	
Data Storage (provide estimate in GigaBytes)	

Access to data across secure NHS network (otherwise known as N3 or HSCN)	Yes / No
Data analysis software (full details to be provided in section 9)	

6.2 Do you want to host a service in the TRE e.g. an application running on a server? Yes / No
If yes, the TRE team will be in touch to discuss your particular technical requirements

6.3 Will your project transfer data into the TRE? Yes / No

If yes, please provide details:

Data Provider(s)	
Data Controller(s) if different from above	
Main contact at data provider (name, email, telephone number)	
Is there a data sharing agreement?	
Has the TRE or any person at CHI been referenced within the project documentation?	
Expected date or frequency of transfer	
Number of files	
Approximate size of each file	
File type	

6.4 Data sensitivity
Referring to document ISMS-07-04 Information Security Classification, state below which classification best matches the data to be imported into the TRE

6.5 Do you need any of the following TRE data management services?

TRE service type	Required?	Details
Access to existing data in the TRE		
Data linkage		
Access to personal data via Secure Data Access Room		
Access for non-UoM project partners		
Access to data across NHS network (otherwise known as N3 or HSCN)		
Validation of dataset		
Support creating metadata		

6.6 Will this dataset be the only copy in existence, or will you be able to download it again?

Only copy / download again

6.7 If you intend to use datasets already in the TRE:

Which existing dataset(s) do you require access to?

Which variables?

Do you require linked datasets (e.g. to consented NHS records) – please include details of legal basis and attach supporting documentation, such as patient information sheet and consent form.

7. Ethical approval

7.1 Have you received/are in the process of obtaining ethical approval? Yes / No

If yes, please provide details of the panel and progress of the application:

Organisation Name	
Address	
Telephone number	
Contact person	
Current application progress	
REC number (if available)	

If no, please specify why ethics approval was not obtained:

Attach copies of evidence (e.g. ethics approval letter) when returning this form. Please reference any documents you are submitting alongside this application form in the field below:

8. Duration of the project

8.1 What is your proposed start date?

8.2 How long do you require the TRE to retain your data?

8.3 How long do you require access to additional services (e.g. analytical software)?

8.4 How do you want your data to be handled after this date?

8.5 Are there any publication deadlines that require derived results to be outputted from the TRE (subject to disclosure controls): Yes / No

If yes, please specify date:

9. Software and Analytical Tools

9.1 Please briefly summarise the tools and techniques you will be using to analyse the data:

9.2 Please list the software you will require to complete your research project and any license requirements you are aware of:

9.3 Will you need additional packages for this software, for example R packages?

10. Research Project members

10.1 Please provide the details of the lead researcher (Principal Investigator):

Name (Title, Name , Surname)	
Institution/organisation	
Institutional email address	

10.2 Please provide details of each Individual, including the PI if necessary, who will require a TRE user account to analyse/process data (expand as required). Each user need to complete training and read TRE user documentation before access can be granted:

Name (Title, name, surname)	Organisation	Email address

11. Data Management Planning and Research Information

11.1 Please provide your Research Data Management Plan record ID: (this can be either the University of Manchester's DMP Tool. Or it can be the DCC's DMPOnline service, or a DMP service local to your Institution, as long as there are UoM personnel referenced in the DMP).

If you do not already have a Data Management Plan, please be aware of the following guidance:

Principle number 2 of the RCUK Common Principles on Data Policy:
http://www.rcuk.ac.uk/research/datapolicy/

And also if the project is storing data at the University of Manchester, principle number 5 of the University of Manchester's Research Data Management Policy:
http://www.library.manchester.ac.uk/using-the-library/staff/research/services/research-data-management/policy/

11.2 If this project already has its funding approved, and it is the University of Manchester that is being awarded the funding, please provide a reference to the project's record on the University of Manchester's CRIS (Pure):

Part B: Declaration by the Principal Investigator

I declare that the information included in this application form and supporting documentation is true and correct to the best of my knowledge.

I understand that any false or misleading information given by me in connection with my application may result in sanctions including termination of the application process or project.

I agree that I will be the main point of contact for updates on the application process and other progress updates.

I agree for my personal information to be processed for the purposes of processing this application and managing the project.

I understand that returning this completed form constitutes an electronic signature.

Name:	
Date:	
Email:	
Telephone number:	

Guidance for validation of datasets as received by data providers

Description of problem

Researchers are often provided with de-identified datasets of routinely-collected data, such as electronic health records (EHR), for research purposes. In many cases a data provider will have to prepare an extract of the dataset which will then be transferred to the researchers. Due the complexity of the datasets, which are often poorly described, the process of preparing a data extract can introduce errors. Such errors could include, for example:

- missing / additional fields
- missing / additional rows
- unexpected format of fields
- unexpected format of files
- unexpected coding of fields (e.g. missing values, dates)
- potentially disclosive information being released (e.g. in free text fields)

Such errors may result in:

- delays in data processing (prepared code is not compatible with the dataset)
- delays resulting from having to request a corrected extract
- knock-on impact on results of analyses, and potentially problems replicating results
- accidental re-identification of patients

These recommendations aim to help minimise the risk of such errors occurring, to encourage validation of datasets to detect potential issues as soon as possible, and to encourage researchers to have a process in place if errors are detected at a later date.

Recommendations

1) Develop a clear data specification in advance of receiving data

Having a clear specification will have three important results: firstly, the researchers should gain an understanding of the datasets that will help them plan their analyses and have realistic expectations about the quality and utility of the data; secondly, it should be useful for the data providers when preparing the extraction; and thirdly, it can be used to validate the data extraction.

Researchers should think about what information will be needed to answer the research question and find out if this is available, in what format, for how many patients and for what date range.

The data specification should be designed by the researchers in close collaboration with the data providers. In the first instance, researchers should request a data specification or data dictionary from the data providers. If this is unavailable, then a new specification will need to be developed. It can also be useful to seek guidance from a clinician/practitioner familiar with the data collection process on the meaning/purpose of the different fields. Questions to ask about potential fields for the specification could be whether the field is entered by users or machine-generated, levels of accuracy, and the meanings of blanks and missing values.

Minimising the risk of re-identification of patients

It is important to consider the potential risk of re-identification of patients within the dataset. Although the datasets should have direct identifiers removed (including names, addresses, exact dates of birth), the richness of the datasets creates the risk of re-identification. Factors that increase the risk of re-identification include: the number of fields requested, the frequency of coded events within fields, and the provision of free text fields. We recommend that researchers:

a) Request the minimum number of fields needed to answer the research question
b) Apply some kind of minimum frequency rule, e.g. do not request codes used less than 10 times in the dataset
c) Do not request free text information unless absolutely necessary. Discuss requirements for free text fields with the data provider as these are particularly risky. There are algorithms for detecting and possibly removing identifiers from free text, and also for converting free text into medical codes. Some are available on the open web, and others are available from CHC researchers such as Goran Nenadic. These could be gathered and used as needed by the research team.

Examples of details to specify in advance

The following list covers some of the key details to agree in advance of a data extract.

a) Exactly what fields are needed, and their names
b) The format of each of the fields requested
c) How missing data is coded within each field
d) How fields such as dates are to be coded
e) Which patients to include in the extract (e.g. provide a list of inclusion/exclusion codes, ages, locations)
f) What date range is required
g) What format the files will be, and the character chosen to delimit the fields

Common sources of errors

These errors are known to have led to problems for data extractions:

The separator used to delimit fields when data is provided in flat text files. Commas are frequently used as separators but can easily appear *within* a data field. This will result in incorrect definition of

fields when the file is read as a table. (This would be apparent if there were more fields than expected in the files). Recommendations:

- Do not request comma-separated files. Instead, request a less frequently used character (e.g. tabs or pipes). Also, advise the data provider to do a search-and-replace within the dataset *before* performing the extraction to remove any instances of your chosen separator within fields.

The presence of carriage returns / line breaks within fields when data is provided in flat text files. This is a particular problem when free text fields are requested: any character may be present within blocks of text. Presence of line breaks within fields will result in additional rows being generated. Recommendations:

- Request the data provider to do a search-and-replace for carriage returns/line breaks within fields before performing the extraction

2) Perform validation checks upon receipt of data

Validation checks should be performed as soon as possible after receipt of the data – it will be easier for the data extractor to resolve any issues in a timely manner, and will ensure errors are detected before the analyses are begun.

a) Ask the data provider for the following details:
   - Number of files
   - Number of fields within each file
   - Number of rows within each file
   - Checksums for each file

b) Compare the files received to the details above. Open each of the files as data tables in your software of choice. Check the number of fields and rows – the incorrect number of fields can imply problems with the delimitation of fields, the incorrect number of rows can indicate erroneous line breaks.

c) Check the format of fields against the data specification

d) Check the data – is it coded as expected (e.g. missing fields, dates)? Simple descriptive statistics will help visualise the data and may highlight unexpected values

e) Carefully review any free text fields for potentially disclosive information. Perhaps take a random sample of the data and review the text.

3) Have a process in place if problems are discovered later

You need a plan for what to do if problems are encountered in any of the validation steps above. Who needs to know within the research group, elsewhere in the institution, and externally e.g. at

the organisation who provided the data? Familiarise the group with information security incident procedure, for example if you find identifying information in a supposedly de-identified dataset.

We collect and use information about you

This helps us to improve stroke care

We take out names and addresses from this information

More than 6000 people in Greater Manchester have a stroke each year

We want to improve their treatment by...

- **being better at recognising** when a stroke has happened
- giving the right treatment **quickly**
- **improving** the **treatment we give**
- **reducing the risk** of another stroke

It will help if we collect and use information in a better way

At the moment there is information

...from **different places** and **different people**

...about what happens in the **ambulance**, from **hospitals**

and **GP records**, and from **brain scans**

We need...

...to link the information with new technology

...to link the information together especially from different health and social care services

We want to see the **whole picture**

what happens to each person after a stroke

what happens across all of **Greater Manchester**

we will then know more about **what needs to improve**

we can **plan for better care** from people with a stroke

Our work at the moment...

...**helping** **paramedics** to identify if someone is having a stroke so they will take people

to the **right hospital**
at the **right time**
for the **right care** for them

...**helping Doctors & Nurses** to use specialist **apps** on phones or tablet computers

The apps will help the **neurosurgery** team to quickly give the **right treatment**

This is important for people who have had a stroke caused by a **bleed in the brain**

An **app** is a software programme
It is usually used with phones, tablets and laptops

...making sure that everyone who has had a stroke has **treatment and support** to **reduce** the **risk** of another stroke

Who can see the information we collect?

Some people already see this information - such as GPs and staff working in stroke care

The **staff** in this project and **people working to improve stroke care** will see it too

Before we use the information -

- **we take out people's names and addresses**

We give **each person a unique number**

No one can link this number with the person

In this way we know **what happens** to people but **not who they are**

www.connectedhealthcities.org/greater-manchester

The website will tell you...

- **who we are and what we do**
- **how we use anonymous patient information**

If you don't want us to use your information
contact the NHS Trust where you received your care

Who is doing this work?

We are part of **Connected Health Cities**

-a project to improve the health of people across the North of England

CONNECTED HEALTH CITIES
A nhsa project

There are **many groups** working together as part of this project

There are **many topics** for this work

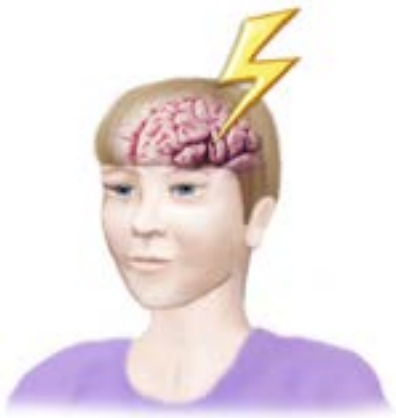
The information in this leaflet is about part of the project

-stroke care in Greater Manchester

Department of Health

The Department of Health pays for this work

What is a stroke?



The **brain needs blood**

Blood keeps the **brain working**

The **blood supply** can **stop**

Then the brain is **damaged**

This can happen suddenly - it is called **a stroke**

It can happen due to

a clot

or a bleed

The **brain** controls **everything we do**

movement

emotions

communication

thinking

A stroke can cause difficulty with some or all of these

Everybody's stroke is different

Data Sharing Agreement
Level 2
GMCHC Stroke Mimics

Salford Royal NHS
NHS Foundation Trust
University Teaching Trust

safe • clean • personal

Classification: Data Sharing Agreement Level 2
Lead Trust Sponsor: Dr. Adrian Parry-Jones
Additional Sponsor: NA
Sponsor's Service: Hyper-Acute Stroke Unit (HASU)
Contact details: adrian.parry-jones@srft.nhs.uk

Third Party Name: University of Manchester
Third Party Contact: Dr. Emily Griffiths, Information Governance Manager
Third Party Service: Health e-Research Centre, Trustworthy Research Environment/Greater Manchester Connected Health City
Care purpose: Service Improvement – analysis of Patients arriving by ambulance to HASU

Data Sharing Purpose:
Analysis to be done in a secure data centre at The University of Manchester as part of service improvement work. The data to be shared are a combination of ambulance and inpatient stroke information for evaluation of "real" and "mimic" patients being brought to SRFT's HASU.

Unique Identifier: Not known
Issue number: New agreement
Replaces: New agreement
Authorised by: Adrian Parry-Jones
Authorisation date: 22/11/2017 *Approved by Jym Bates 26/6/18*
Next review: Normally 2 years from date of issue
Distribution List:

1 Introduction

This document provides a means of establishing a standard for the sharing of information in respect to SRFT's HASU and is intended to form the basis of a model of good practice for information sharing between the organisations listed in Section 11 in compliance with the Data Protection Act 1998 and the Caldicott Principles.

This agreement covers the sharing of information for any of the purposes listed in section 3.1 and comprises the common principles and procedures which will be adopted wherever and whenever these organisations have to share information for these purposes.

This agreement should be read in conjunction with the overarching level 1 agreement.

This agreement is intended to cover the following types of data sharing: -

- Non Personal Data. Information that does not relate to people; e.g. information about organisations, natural resources and projects, or information about people that has been aggregated to a level that is not about individuals.
- De-Personalised Data. Information that relates to individuals, but where it is not possible to identify individuals from the information, whether in isolation or in conjunction with other information that the organisation holds.
- Personal (Sensitive) Data. Information that relates to individuals where the individual can be identified from the data and also where the purpose of the sharing is for research purposes, including statistical or historical purposes. (Only the third situation falls within the remit of the Data Protection Act 1998 and benefits from a special exemption (Section 33 of the Act) which allows data to be used for research even if it was not collected for this purpose. Personal data held only for research purposes may also be kept indefinitely. Other data sharing situations (i.e. sharing of personal data for other than research purposes) should also be reviewed under Caldicott Principles and the Pseudonymisation Implementation Project. This includes statutory obligations to share data; the appropriate statutory authority should be explained in section 3.1 of the agreement.

Where the agreement is for personal (sensitive) information all parties must be registered with the Information Commissioner and have relevant purposes specified in their scope of registration. Evidence of this will be demonstrated by writing the organisation's registration number in the appropriate boxes of Section 3.4. If there is any doubt about a partners scope of registration the other party(ies) should satisfy themselves on this point by checking the online public register of data controllers: -
https://ico.org.uk/about-the-ico/what-we-do/register-of-data-controllers/

2 Aims

This agreement provides a framework for the secure and confidential sharing of information between organisations to: -

- Ensure service users / patients receive the health / social care services they require
- Provide seamless and coordinated care
- Work effectively and efficiently together to tailor services to the particular circumstances and requirements of each individual
- Meet the needs for communities and individuals for care, protection and support
- Set out for service users / patients the reasons why information about them may need to be shared and how this sharing will be managed and controlled so that confidentiality is maintained

3 Objectives

3.1 Data Sharing Purpose Statement

For the purpose of the HASU service improvement as part of the GM stroke ODN, information in the form of de-personalised data will be shared to find out how often the 'Yes' decision is correct at SRFT's HASU, and provide feedback to reduce inappropriate arrivals: -

- Analyse the outcomes of patients brought by ambulance to SRFT's HASU from 1st September 2015 to 28th February 2017, i.e. whether they were diagnosed with a stroke or not.
- Compare the features and symptoms of patients who were diagnosed with those who had a stroke "mimic".
- Seek to update guidance to paramedics on what are and are not signs of a stroke, with a view to reducing the proportion of HASU arrivals of patients not requiring acute stroke care.

- Data from HASU care and digitised from NWAS records will be sent from Salford's EPR to the University of Manchester (UoM; see Figure 2 below).
    - Previous communications with NWAS determined that they do not have electronic records, nor are likely to in the immediate future. To inform and deliver service improvements, Greater Manchester Connected Health City at the University of Manchester is paying for staff at SRFT to digitise the NWAS sheets for HASU patients. The cohort All these staff members have signed relevant agreements for access to SRFT computer systems. The cohort of patients has been determined by an analyst at SRFT in conjunction with the PI. The de-personalised, digitised records from Salford will inform HASU service improvements and advice to NWAS.

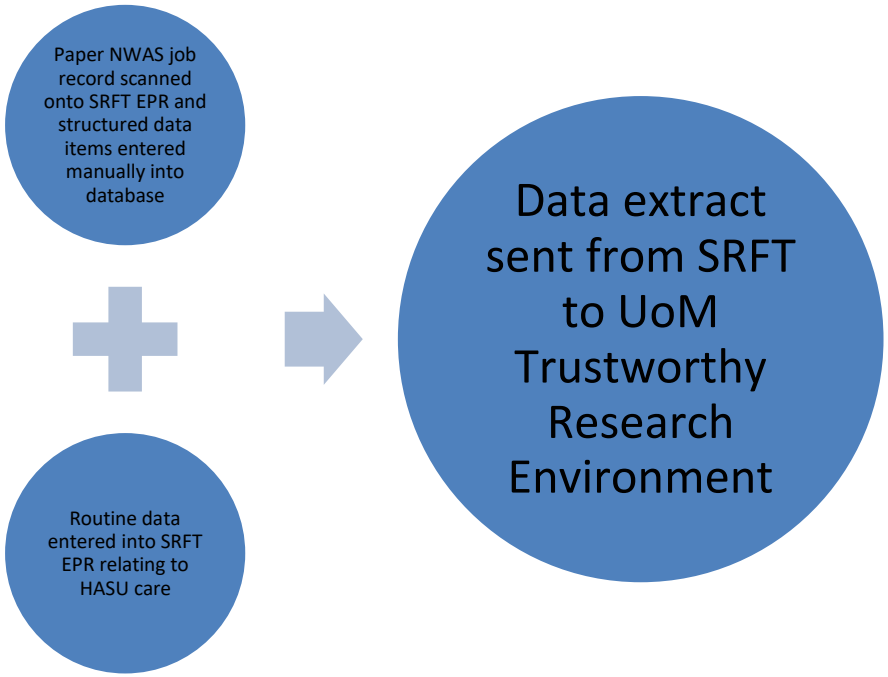


Figure 1 – Proposed data sharing from SRFT to UoM. The NWAS data come from sections 1,2,3, and 4 of the RX7 form. Identifiers will not be shared (i.e. staff number, patient NHS number, patient first or last name, patient address, patient date of birth, patient GP, and next of kin name, contact number, or any free text fields). Postcodes will be aggregated according to a routine algorithm by analyst at SRFT before transfer of data. A more detailed data flow and stroke pathway diagram is included in the accompanying PIA, and a full data specification is enclosed.

3.2 Data Ownership

SRFT remains the owner of the data, and grants to UoM a license to analyse the data (see level 1 agreement).

3.3 Conditions on Use of Supplied Data

Only the following named individuals (all UoM employees) will have access to the data, and only for the purposes outlined above:

- Camilla Sammut-Powell
- Adrian Parry-Jones

Data will be accessed securely under strict information governance and information security rules in line with ISO27001 and IG Toolkit requirements.
Acknowledgement will be granted in any publications or presentation of results
No fees are payable.
For further details see level 1 agreement.

3.4 Conditions on Use of Resulting Data

If UoM seeks to amend the purpose of data use above, permission will be sought beforehand in writing to SRFT.
SRFT will be consulted prior to publication, and has the right to see results at any time on request.
Generic metadata may be published on UoM website in line with UoM library and HeRC's policies. All such listing will follow FAIR principles.

3.5 Measures to Ensure Security of Data

Data will be handled and deleted securely under strict information governance and information security rules in line with ISO27001, IG Toolkit requirements, and the standard operating procedures of the Trustworthy Research Environment.

3.6 Retention Period for Supplied Data

Data will be retained for the length of the level 1 agreement, and longer if required by any journals in which results are published.

3.7 Format of Supplied Data

Secure electronic transfer of delimited text files via FTP across the internet (preferably N3 if possible) from SRFT server into the Trustworthy Research Environment.

3.8 Other Conditions

Both organisations to share information about these data to aid valuation of the stroke service improvement work.

4 Legislation & Guidance

Sign-off of this agreement signifies that all constituent parties, fully comply with the Data Protection Act 1998, together with all other related and relevant legislation and guidance covering issues of data collection, sharing, transmission and storage, including:

- The NHS Information Security Code of Practice 2007;
- Confidentiality: NHS Code of Practice, 2003;
- The Caldicott Report, 1997;
- The Freedom of Information Act, 2000;
- NHS Records Management Code of Practice, 2006 & 2009;
- NHS Care Record Guarantee, 2007;
- Social Care Record Guarantee, 2007;
- Electronic Communications Act, 2000;
- Health and Social Care Act 2012
  Health and Social Care (Safety and Quality) Act 2015
- The Re-Use of Public Sector Information Regulations, 2005.

Any organisation processing NHS data is required to comply with the NHS Information Governance Toolkit. To aid transparency it is recommended that constituent parties undertake a review of information governance using the NHS Information Governance Toolkit. This will ensure any relevant outstanding issues relating to information governance are identified.

5 Information Sharing Principles

In seeking to share information organisations will adhere to the following principles: -

- The organisations that are party to this agreement are committed to enable data to be shared in a manner that is compliant with their statutory responsibilities
- Service users / patients and carers will be fully informed about information that is recorded about them and as a general rule, be asked for consent before information is shared with colleagues or another organisation. This consent should be clearly recorded.
- The rules regarding disclosure of information apply to service users who lack capacity to consent. Where appropriate consent should be obtained from the person with the legal authority to act on the person's behalf. The reasons for the final decision should be clearly recorded.
- Organisations will ensure that staff receive appropriate training around service users / patient confidentiality

- Where professionals request that information supplied by them be kept confidential from the people who use services, the outcome of this request and the reasons for taking the decision will be recorded
- Information will not be used for any other purposes or further shared without prior consent of the user

6 Commitment to Developing Standard Procedures

The adoption of standard operational procedures should govern the exchange of information between the organisations that are party to this agreement. These processes will set the standards that the organisations will be expected to work towards and should be common to all service specific operational information sharing agreements.

7 Formal Approval / Adoption

This agreement applies to all organisations detailed in Section 11. It also applies to all staff, temporary and volunteer workers within the organisations that are party to this agreement.

8 Dissemination / Circulation of the Agreement

Staff will be made aware of and have access to this agreement. Any staff guidance will be read in conjunction with this agreement.

The agreement will be communicated to service users / patients, carers and voluntary organisations to ensure that individual rights in relation to the disclosure of personal information are upheld.

9 Data Subject Access Requests

Each partner organisation will designate an appropriate manager with the authority to make decisions with regard to subject access requests. All such request and any actions taken must be properly recorded within the partner organisations management systems.

10 Complaints

Each partner organisation will designate an appropriate manager with the authority to make decisions with regard to complaints. All such requests and actions taken must be properly recorded within the partner organisations management systems.

11 Organisations and Signatories

Organisation Name	Salford Royal NHS Foundation Trust
Signatory (inc Title)	*Approved by Jym Bates 26/6/18*
Signature	
Date	
Organisation Name	University of Manchester
Signatory (inc Title)	Prof. Niels Peek Head of Greater Manchester Connected Health City
Signature	
Date	

APPENDIX 1 - INFORMATION FLOWS:

Flow number/ code	**Sender organisation and department**	**Recipient organisation and department**	**Purpose / Legal Basis e**.g. Care & Treatment	**Data Items***	**Method of transfer /format/ applicable data standards**	**Frequency of data sharing/ access**	**Retention period**	**Justified by and date**
Not known	Information Services, SRFT	Trustworthy Research Environment, Health e-Research Centre, University of Manchester	Anonymised data to inform stroke service improvement work	De-identified NWAS data digitised from RW7 form and HASU data from EPR. *See data specification.*	sFTP	Transfer of retrospective data from 1st Sept 2015 to 28th February 2017. We anticipate two transfers in early 2018, each covering around half of the digitised data.	5 years	*Approved by Jym Bates 26/6/18*

CONFIDENTIAL

DATA TRANSFER AGREEMENT

THIS AGREEMENT is made the by and between

The University of Manchester (acting in this instance through the Greater Manchester Connected Health City) with a business address at Oxford Road, Manchester, M13 9PL, United Kingdom ("Receiving Organisation");

And

Pennine MSK Partnership Ltd, with an address of Integrated Care Centre, New Radcliffe Street, Oldham, OL1 1NL United Kingdom ;
And

Salford Royal NHS Foundation Trust, with an address of Stott Ln, Salford M6 8HD United Kingdom ("the Providing Organisation");

Pennine MSK Partnership Ltd and Salford Royal NHS Foundation Trust referred to as "the Providing Organisation" as applicable,

each a "Party" and collectively the "Parties"

WHEREAS Receiving Organisation is the base and the legal entity for the Greater Manchester Connected Health Cities Hub and holds anonymised data in the field of Healthcare; and

WHEREAS the Providing Organisation collects certain Data which shall be detailed in the in the form annexed hereto at Schedule 1.

WHEREAS The Parties agree that before each transfer of Data a Schedule 1shall be signed by the Parties and shall detail the type of Data required, the reason for the required Data/the project and the use of such Data.

WHEREAS the Parties wish to clarify their respective rights and obligations in respect of the Receiving Organisation's use of the Data and each Party's use of the results of such project through entry into this Agreement.

NOW THEREFORE in consideration of the mutual promises and covenants set forth herein, and intending to be legally bound, the Parties agree as follows:

1. Definitions

1.1 **"Data"** shall mean the data described in the Schedule 1 together with any additional information made available relating thereto.

1.2 **"Purpose"** shall mean use by the Receiving Organisation for purposes in the specific project detailed in this Agreement and no other purpose.

1.3 **"Receiving Organisation's representative"** shall mean the representative(s) of the Receiving Organisation named in Schedule 1 responsible for the conduct of the Purpose.

1.4 **"Results"** shall mean the results arising Purpose performed by the Receiving Organisation using the Data including without limitation all analyses, calculations, algorithms and meta-data irrespective of format.

1.5 **"Schedule 1"** means the schedule attached to this Agreement between the Parties which is a template detailing the particular Data transfers which must be signed by both Parties before any Data is transferred.

2. Transfer of Data

2.1 Prior to each transfer of Data the Parties shall both sign a Schedule 1 which shall set out the details of the transfer such as the type of Data required, the start and end date for the use of such Data and the purpose of the transfer. For the avoidance of doubt, no transfer of Data shall take place until such Schedule 1 has been signed.

2.2 The Providing Organisation shall provide the Data described in the Schedule 1 to the Receiving Organisation as soon as reasonably possible following execution of the particular Schedule 1. Data will be transferred to the Receiving Organisation's representative by the Providing Organisation's representative responsible for the care of the Data. The Data is supplied by the Providing Organisation in pseudonymised form without the Psuedonymisation key or other means for the Receiving Organisation to re-identity individuals from the Data.

2.3 The transfer of the Data under a Schedule 1 shall be governed by the terms and conditions of this Agreement.

3. Treatment of Data by the Receiving Organisation

3.1 The Receiving Organisation's representative will be responsible for receipt of the Data and upholding any obligations on the Providing Organisation in respect of the Data. Notwithstanding the foregoing, the Receiving Organisation agrees to maintain Data disclosed or transferred to the Receiving Organisation by or on behalf of Providing Organisation as confidential data with the same degree of care it holds its own confidential data.

3.2 The Receiving Organisation will mark and store the Data in such a manner that it is at all times traceable as proprietary to the Providing Organisation. The Receiving Organisation will keep the Data secure using password protection as a minimum and preferably data encryption. The Receiving Organisation will not store the Data on a laptop, disc, external drive or any other temporary media.

3.3 The Receiving Organisation will not use the Data or cause the same to be used except for the Purpose. The Receiving Organisation will disclose such Data only to its directors, officers, employees, faculty, and researchers directly concerned with carrying out the Purpose subject to the Receiving Organisation having in place with

such persons obligations no less strict than those set out herein and remaining fully liable for any breach by such persons. The Receiving Organisation will neither disclose the Data to any third party nor use such Data for any other purpose without the prior written consent of the Providing Organisation.

3.4 Each Party shall comply with the Data Protection Act 2018 ("the 2018 Act") and any other applicable data protection legislation or regulations. The Parties agree to use all reasonable efforts to assist each other to comply with the 2018 Act. For the avoidance of doubt, this includes providing the other with reasonable assistance in complying with subject access requests and consulting with the other prior to the disclosure of any personal data created in connection with the conduct or performance of the Purpose in relation to such requests.

4. Exceptions.

4.1 The Receiving Organisation's obligations of nondisclosure and the limitations upon its right to use the Data shall not apply to the extent that the Receiving Organisation can demonstrate that Data: (a) was in its possession prior to the time of disclosure without obligation to the Providing Organisation; or (b) is or becomes public knowledge through no fault or omission of the Receiving Organisation; or (c) is obtained by the Receiving Organisation from a third party under no obligation of confidentiality to the Providing Organisation; or (d) if the Receiving Organisation is requested or ordered to disclose Data in connection with a legal or administrative proceeding, the Receiving Organisation will give the Providing Organisation prompt notice of such request. The Providing Organisation may seek an appropriate protective order or other remedy or waive compliance with the provisions of this Agreement or both. If the Providing Organisation seeks a protective order or other remedy, the Receiving Organisation will cooperate with the Providing Organisation, at the Providing Organisation's expense. To the extent the Providing Organisation fails to obtain a protective order or waive compliance with the relevant provisions of this Agreement, the Receiving Organisation will disclose only that portion of the Data which its legal counsel determines the Receiving Organisation is required to disclose.

5. Ownership of Results and License.

5.1 Nothing shall affect the ownership of the Data which shall remain vested in the Providing Organisation. For the avoidance of doubt this extends to any of the Data which is included in the Results.

5.2 Results of the Purpose will be owned by the Receiving Organisation. The Receiving Organisation will keep the Providing Organisation reasonably updated with progress of each project and will supply all Results to the Providing Organisation in a timely manner.

5.1 The Providing Organisation will be entitled to receive all raw data contained in the Results. The Receiving Organisation grants to the Providing Organisation, without fee or payment of any kind, a perpetual, worldwide, non-exclusive, fully sub-licensable, license to use the Results for internal teaching and non-commercial research.

6. Publication.

6.1 Notwithstanding the other provisions of this Agreement, it is the desire of both Parties for Results to be published as appropriate.

6.2 The Receiving Organisation may publish the Results, will detail the source(s) of data used, and to the extent that the Providing Organisation has made any significant contribution to the Results the Receiving Organisation will give due acknowledgement to the Providing Organisation which shall be agreed at the time of publication in accordance with accepted scientific publication practice.

6.3 The Providing Organisation shall have the right to review all proposed publications which refer in any manner to the Results or the Data. The Receiving Organisation will send relevant manuscripts and proposed publications to the Providing Organisation for review prior to publication. The Providing Organisation will have up to thirty (30) days to review each manuscript ("Review Period") and will have the right to delete any Providing Organisation confidential information or Data from such manuscript. The Receiving Organisation will be free to publish at the end of the Review Period.

7. Retention of Rights in the Data.

7.1 All intellectual property rights in the Data shall remain the property of the Providing Organisation at all times. Nothing in this Agreement shall be construed as granting any license to the Data or the Providing Organisation's other intellectual property rights unless otherwise expressly set out under this Agreement or the Schedule 1.

8. No Further Obligation.

8.1 Providing Organisation is disclosing Data to Receiving Organisation on the express understanding that neither Party will be obligated to enter into any further agreements relating to the subject matter hereof, and unless and until any final definitive agreement with respect to the above subject matter is agreed between the Parties, the Parties will not have any obligation to the other Party except under this Agreement or any other definitive written agreement already entered into with respect to the subject matter.

9. Limited Representations and Warranties.

9.1 The Providing Organisation represents that the Data has been derived and supplied in accordance with all applicable laws, rules and regulations.

9.2 Each Party represents to the other, to the extent that it supplies to the other Party, uses itself, or permits the other Party to use Data it has obtained from a third party to perform this Agreement that it has all required permissions, licenses and consents from such third party to do so.

9.3 The Receiving Organisation warrants it will conduct the permitted project using the Data and any third party data in accordance with all applicable laws, rules and regulations; and save for the express warranties set forth in this Agreement, no representations, undertakings or warranties, whether express or implied, are made or given by either Party including without limitation (i) as to the accuracy, completeness, or fitness for a particular purpose of the Data; or (ii) the Results or their freedom from infringement of any third party intellectual property rights.

10. Term and Termination.

10.1 This Agreement will terminate five (5) years from the date of last signature.

10.2 Either Party may terminate this Agreement without cause on giving the other not less than thirty (30) days written notice. Those sections intending to survive expiry or earlier termination of this Agreement will survive.

10.3 For the avoidance of doubt, if the period of a Schedule 1 extends beyond the termination date of this Agreement, the terms and conditions of this Agreement shall remain in force in relation to the data transfer under the Schedule 1until the expiry of the Schedule 1.

11. Return or destruction of Data.

11.1 Data shall be returned or destroyed as set out in the Schedule 1. Upon the written request of the Providing Organisation at any time or following expiry of the Schedule 1, the Receiving Organisation will promptly return or destroy (as set out in the Schedule 1) all Data in its possession or control and all copies of it, save that the Receiving Organisation will not be required to surrender or destroy any computer files stored securely by the Receiving Organisation, its business units and affiliates that are created during automatic system back-up or retained for legal purposes by the legal division of the Receiving Organisation. The Receiving Organisation will certify to the Providing Organisation that all electronic copies other than those required as above have been deleted, and that all paper copies have been destroyed.

12. General

11.1 **Headings.** Headings are provided for convenience only and do not affect the construction or interpretation of this Agreement and the Schedule 1.

11.2 **No Waiver.** No waiver shall be binding unless in writing and signed by the Party making such waiver. A waiver made on one occasion shall not be deemed a waiver on any other or subsequent occasion. All rights of the Parties are cumulative.

11.3 **Authority.** Each Party represents to the other that it has the full authority to enter into this Agreement and the Schedule 1. Each Party represents to the other that it is entering into this Agreement as principal not agent. Each signatory represents that they have the full authority to bind their respective company or organization to the terms of this Agreement.

11.4 **Entire Agreement.** This Agreement and the Schedule 1 sets forth the entire agreement between the Parties as to its subject matter and supersedes all prior discussions, understandings, or verbal agreements (if any) in relation thereto all of which are replaced in their entirety by the terms of this Agreement. Notwithstanding the foregoing, this Agreement and any Schedule 1 shall not supersede or vary any other definitive written agreements already executed between the Parties. For the avoidance of doubt, in the event of conflict between this Agreement or any Schedule 1, the terms of the Agreement shall prevail.

11.5 **Variation.** None of the terms of this Agreement shall be amended except in a writing signed by each Party.

Schedule 1

Schedule 1 Template

Document Management

Revision History

Version	Date	Summary of Changes
1.0	01/08/2016	First draft for internal review
2.0	20/03/2017	Incorporating minor changes following stakeholder review
3.0	13/07/2017	Amended to go under University of Manchester Agreement

Reviewers

This document must be reviewed by the following people:

Reviewer name	Title / Responsibility	Date	Version
CHC IG working Group			

Approved by

This document must be approved by the following people:

Name	Signature	Title	Date	Version
John Ainsworth			03/05/2017	1.0

1. Data Sharing Agreement

This Schedule 1 is subject to the terms and conditions set out in the Agreement above.

1. Title and Reference Code

CHC Project	Musculoskeletal disease data hosting
Reference	

2. Parties to the Agreement

Agreement owner	The University of Manchester
Receiving Organisation	The University of Manchester
Receiving Organisation's Representative	Greater Manchester Connected Health City
Providing Organisations	Pennine MSK Partnership Salford Royal NHS Foundation Trust
Providing Organisation's Representative	Not applicable

3. Term of the Agreement

Start Date	01/09/2018
End Date	31/07/2020

4. Privacy Impact Assessment

Summary of Privacy Impact Assessment for Data Sharing	The Privacy Impact Assessment statement found residual risk to be low or medium. A communications plan is being developed which will further mitigate risk.

5. Data Details

Type of Data to be Shared Anonymised (inc. Pseudonymised/ De-identified)	Pseudonymised MSK patient records
Purpose for Sharing	Research project into the burden of musculoskeletal disease in secondary care (Principal Investigator Jenny Humphreys)...
Personnel to have access to the Data	Research project data analysts: Jenny Humphreys
Details of the Data to be shared	• Pseudonymised id • Age at that visit • Gender • IMD score • Postcode (first 4 characters) • Ethnicity (collapsed) • Date of appointment • Diagnosis list • Any coded onward referrals to other services from that visit
Details of how the Data will be shared	Material transfer via secure FTP from the Providing Organisations to the Trustworthy Research Environment at The University of Manchester. Only named users above will access the data, which will be held

Case Study A

1. A - Introduction

A study investigating antimicrobial resistance seeking prescribing information from GP practices to assess the quality of treatment decisions for patients with acute bacterial infections.

2. A - Scoping and landscaping

We started with a focus on the problem: rising antibiotic resistance and the problem of overprescribing, often by GPs. We arranged an event to pull together key stakeholders regarding quality prescribing in general practice. At this event, stakeholders pulled ideas for what they would want in a data-driven tool to inform them about how to make improvements.

The researchers had experience using GP data available nationally. This preliminary analysis informed the specification of data sought from local providers. Data were sought from the new local integrated care record as this would include data from general practice and from patients admitted to hospital experiencing complications from antibiotic resistance.

3. A - Approvals

Although the local integrated care record had planned to make linked data available for research, the systems were not in place in time for our project. We decided instead to drop the hospital element and just get data directly from pilot practices. This involved getting Health Research Authority (HRA) approval to conduct the research study with a particular IT provider who could extract data from participating GP practice computer systems. The data would be sent to the university for analysis by the project analysts. The university signed a contract with this IT company. To participate practices needed to have their Caldicott Guardian sign an agreement with the university and another with the IT company.

4. A - Project planning

From the initial draft of an application form, HRA approval took 7 months. Agreements with the IT provider were drawn up after this in the spring and needed a minor amendment (changing from the integrated record). The system to process the data took a further 6 months to build, with added time for checking and approving the data request (meeting the data specification proved more tricky than initially envisaged). From drafting the initial data specification to showing the analytical tool to participating practices took over 18 months.

The delays in getting permissions and especially in building the infrastructure meant that recruitment of practices needed to stop. There was also only a 2 month window to deliver the solution to practices that had already signed up.

Case Study B

1. B - Introduction

A study into a chronic respiratory disease (COPD) investigating pathways of inpatients and outpatients across a conurbation.

2. B - Scoping and landscaping

A planning meeting was held between an existing member of the project and a new manager, which identified NHS Digital as the main source of data across various healthcare providers in the region. The application form was completed and went via the Data Services for Commissioners Regional Office (DSCRO), which is a local office that deals with NHS Digital data.

3. B - Approvals

The application underwent a two-step approvals process, and in between we had a face to face meeting with a representative of NHS Digital. They required payment of a five-figure sum. There were some delays in getting the data. Initially we were told the data would flow within 24 hours of board approval but in the end it took over 3 months, and in the meantime we amended our request to cover a slightly larger geographical footprint and to extend the time that we could access the data.

4. B - Project planning

It took 7 months from initial, serious enquiry with NHS Digital to receiving data from them. In this timeframe new data protection regulations (the GDPR) came into force which resulted in additional queries that came up regarding the status of different organisations some of which needed to renew their data controller registration with the Information Commissioner’s Office (ICO).

Case Study C

1. C - Introduction

A study aiming to make more health information about vulnerable families available to social workers.

2. C - Scoping and landscaping

Our project aimed to share data among multidisciplinary Early Help social care teams to support the care they offer to families. The project team is made up of university health services researchers and the technical team at the university and a local company. There was also a qualitative researcher who interviewed families and practitioners who are involved with Early Help services about their experiences of information sharing.

We participated in a steering group across five local authorities, project members, and some other researchers. Liaison was needed with a litigation manager, an information rights officer, clinical safety officers, and care professionals. The steering group determined that a new Interoperability system would supply data, and that the project would be hosted securely at a private IT company.

3. C - Approvals

The qualitative researcher needed university research ethics committee approval before seeking interviewees. The project team needed to build the system to comply with clinical safety requirements. Information sharing agreements needed to be in place for each participating care provider and with each Local Authority. Once the agreements were signed there was a 6-8 week timeframe before data were received.

4. C - Project planning

Project management was led by a health informatics service team. Building a network of enthusiastic stakeholders and getting signed agreements has been time consuming. Following delays we had to focus on the data flowing to one local care team as a proof of concept involving just one local GP practice as a service improvement pilot.

At the start of the project we were aware that information governance would be time consuming. One member of the project left early, and another went on parental leave. Remaining members of the project team has goodwill but no specialist experience in information governance, and the new data protection legislation (the GDPR) posed uncertainty. In hindsight it would have been beneficial to have had more representatives from across health and social care involved in the steering group.

Case Study D

1. D - Introduction

A study aiming to understand complex health problems in the very elderly population across one city using statistical modelling.

2. D - Scoping and landscaping

We started with a focus on the problem: Population modelling work investigating frailty and multimorbidity to improve health and wellbeing of a city's population. We sought linked, pseudonymised data including primary, secondary, and adult social care data for the city.

3. D - Approvals

We needed to ensure registrations and approvals were in place with the Information Commissione''s Office, local medical committee, senior stakeholders (including Caldicott Guardians) within Trusts and Local Authorities, the Clinical Commissioning Group (CCG) forum, and regional Commissioning Support Unit (CSU).

Paperwork we needed to complete included Data Protection Impact Assessments (DPIAs), and data sharing agreements with each data provider.

4. D - Project planning

It took six months to get HRA approval to approach GP practices, then to sign agreements with 88 GPs, 3 trusts and one local authority.

Once approvals were in place we needed to deal with technicalities of extracting data from organisations, for example the CSU provided 'SUS' data about users of secondary care. Some of this involved special "black box" software from a company that could extract data from GP systems. The networking and firewalls also needed arranging with BT. This all took 18 months from agreements being signed to getting data from primary care. During this time period we worked on pilot projects with some GP sites to try to get things moving.