# Request for Data

<table>
<thead>
<tr>
<th>Date of Submission:</th>
<th>30th June 2017</th>
</tr>
</thead>
</table>

**Submitted by and job role:** Dr Adrian Parry-Jones  
NIHR Clinician Scientist/ Honorary Consultant Neurologist. SRFT & Greater Manchester Connected Health Cities.

**Reviewed by Quality Team at NWEH:** Not applicable

<table>
<thead>
<tr>
<th>Comments:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Approved by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marie Kane</td>
</tr>
</tbody>
</table>

Send completed form to [datarequest@srft.nhs.uk](mailto:datarequest@srft.nhs.uk)
Summary of Project: To identify ways of improving the quality of secondary prevention for stroke and TIA patients following discharge from the acute setting.

(High level description): This project aims to improve the quality of care across the stroke pathway, from suspected stroke onset through to long-term care, by linking and analysing routinely available datasets collected across different settings and services. This will help us to better understand the patient journey, to identify opportunities for improvement, and to design, implement and evaluate tests of change in a number of different areas identified for improvement whilst also establishing a ‘learning health system’ across Greater Manchester.

The work of the project will be broken down into four work streams corresponding to four parts of the stroke pathway. The purpose of this SIR request (#243) is to support the work stream which aims to improve the quality of secondary prevention for stroke and TIA patients following discharge from the acute setting, with a particular focus on atrial fibrillation detection and treatment, and blood pressure measurement and management.

We propose to use the SIR data to provide linked historical Primary Care GP data and acute secondary care data to create a cohort of thousands of stroke and TIA patients. These data will be de-identified and shared with the university to examine current practice (including variation in practice) with regards to stroke secondary prevention, and explore the reasons why secondary prevention may currently be suboptimal. These data will also be used to develop a model to predict those patients who are at highest risk of recurrent stroke/TIA to support service improvement efforts.

What is/are the research questions?

<table>
<thead>
<tr>
<th>Primary</th>
<th>1a - Atrial fibrillation detection:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. What percentage of the study population with ischaemic stroke/TIA has known AF? Of these, what percentage was taking antiplatelet drugs or anticoagulants on admission?</td>
</tr>
<tr>
<td></td>
<td>2. What percentage of the study population with ischaemic stroke/TIA and without previously known AF has a 12-lead ECG performed and reviewed during their admission?</td>
</tr>
<tr>
<td></td>
<td>3. What percentage of the study population with ischaemic stroke/TIA, without previously known AF, and with sinus rhythm on their hospital ECG goes on to have prolonged ECG recording? What percentage of these identifies new AF?</td>
</tr>
<tr>
<td></td>
<td>4. What percentage of the study population with ischaemic stroke/TIA and without previously known AF have 12-lead ECGs</td>
</tr>
</tbody>
</table>

Send completed form to datarequest@srft.nhs.uk
in primary care following admission? What percentage of these identifies new AF?

1b – Atrial fibrillation treatment:
1. Following each of these opportunities to treat (premorbid AF, hospital ECG shows AF, prolonged recording shows AF, primary care ECG shows AF) what proportion of patients commenced (or restarted) an antiplatelet or anticoagulant drug?
2. What was the delay between identification of AF and commencement of treatment?
3. Is there documentation of a risk-benefit discussion with the patient (e.g. risk scores recorded)?
4. What proportion of patients with ICH and premorbid AF commence or restart anticoagulation/antiplatelet drugs after the index ICH?

2a: Blood pressure measurement:
1. How often (readings/day) is blood pressure recorded during the hospital admission and what is the mean and SD of these readings during the admission?
2. How often (readings/month) is blood pressure recorded after the hospital admission, who measures it, and what is the mean and SD of these readings?

2b: Blood pressure treatment:
1. When are changes in antihypertensive drugs made, what are these changes and who initiates them?
2. How compliant are patients with their antihypertensive drugs (*using standard measures derived from frequency of repeat prescriptions*)?

3: Recurrent strokes:
1. In the whole study population, what proportion of patients has a recurrent stroke (considering ischaemic stroke and ICH separately)?
2. After adjusting for other factors associated with risk of recurrent ischaemic stroke, are the following factors associated with risk of ischaemic stroke?
   a. Anticoagulation (*we may wish also consider time taken to start*)
   b. mean SBP & DBP after index event
   c. SD of SBP & DBP after index event
3. After adjusting for other factors associated with risk of ICH, are the following factors associated with risk of ICH?
   a. Anticoagulation (*we may wish also consider time taken to start*)
   b. mean SBP & DBP after index event

Send completed form to datarequest@srft.nhs.uk
What data will be required to answer this question:

| How will the cohort of patients under investigation be defined? | Patients registered with a GP in Salford who have a Read code to indicate a diagnosis of stroke/TIA with an input date since the start of April 2007. The stroke and TIA read codes that will be included are available from NHS Digital (pdf download; see pages 15 and 17). We understand the first cohort (Apr 2007-June 2017) to comprise about 6,500 patients. To keep analyses up to date and track service improvements we are also requesting an update every 3 months containing data for any patients or events not sent previously. |
| Do you want all the journal data for patients in the cohort to be extracted? If so, please justify | Yes. Other medications, diagnoses and risk factors may be associated with risk of stroke and we need to check this as per the research questions above, such as Q3-2. Our primary aim with this work is to identify where care is sub-optimal in terms of stroke secondary prevention, and to then identify what factors are associated with this so that targeted improvement interventions can be made and their impact subsequently tracked using a prospective data feeds. To obtain as detailed an understanding as possible and to identify unexpected predictors, we will require all journal data. Variables such as sex, ethnicity, and LSOA are commonly used in health research as indicators of deprivation. Analysing a wide range of covariates will ensure any resulting service improvements are well-targeted and have a higher chance of reducing recurrent stroke in future Salford patients. |
| If not specify how the data should be restricted | In order to fully de-identify the dataset while also keeping it adequate for the research needs, the following variables used in processing the data have been excluded from the dataset that will be transferred:
  - NHS number [for data linkage at SRFT]
  - Postcode
  - Date of Birth
  - Date of Death
  - Date of onset and associated dates
  - GP practice code

  In some cases, coarser alternatives have been chosen:
  - Patient pseudonym
  - GP practice pseudonym
  - The Lower Layer Super Output Area (LSOA) geographic identifier derived from the patient’s postcode
  - Year of birth |
• Month and year of death (mm/yyyy)
• Month and year of date of onset and
• For all other dates the number of days relative to date of onset, e.g. discharged 7 days later

A full list of variables and tables in the final dataset is included below.

Is secondary care data required? If so, specify

Yes. Primary care data is to be linked to data from the Salford Royal NHS FT Electronic Patient Record to create a comprehensive dataset that follows patients’ stroke ‘secondary prevention journey’.

The following data from EPR pertaining to the cohort of patients under investigation is required:

• Clinical observations
• Medication
• Health issues
• Diagnosis
• Investigations

Data will also be linked to that contained within the SRFT SSNAP database.

A full list of variables and tables in the final dataset is included below.

I accept that the data will be supplied in anonymised form i.e. names, addresses, postcodes will be removed and NHS number replaced with an anonymised patient ID.

Who will have access to the data?

Name & Job Title

Two analysts employed in the Faculty of Biology, Medicine and Health at The University of Manchester who work on Greater Manchester Connected Health Cities projects. Both have signed honorary contracts with Salford Royal NHS Foundation Trust:

Dr. Matthew Sperrin, Senior Lecturer in Health Data Science

Dr. Camilla Sammut-Powell, Medical Statistician and Research Associate, Greater Manchester Connected Health Cities

Where will the data be held?

Trusted Research Environment,
Health e-Research Centre,
Vaughan House,

Send completed form to datarequest@srt.nhs.uk
How will the results of this study be disseminated?

| What data will be shared? | Baseline summary data (e.g. means, standard deviations, counts, proportions).  
Correlations, associations and other inferential statistics derived from the data.  
We will never share small cell count data (<5 subjects). |
|--------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| With Whom & How          | With the wider scientific community through journal papers, reports, and conference presentations.  
Any press/media interest will be co-ordinated jointly by The University of Manchester (on behalf of Connected Health Cities) and SRFT. |

How long will the data need to be stored?

The full dataset will be retained in the Trusted Research Environment for the duration of the project (Connected Health Cities is funded until 31/03/2019) and, if any publication based on it has requirements in terms of data storage for replication of results, the final dataset used for analysis will be retained in the university’s secure file vaults for as long as required by the journal.

When will the data be deleted?

| How will it be deleted | Even though the data are de-identified, The University of Manchester holds good research conduct and duties of confidentiality very seriously. Data will not be stored on individual PCs or any removable media. Data will be stored securely on the Trusted Research Environment servers within the data safehaven, for which ISO27001 accreditation is being sought this year. Data will be deleted securely once the project ends and sent securely to University of Manchester secure file vaults to be retained for the number of years required by any academic journals in which the project findings are published.  
If any server hardware is to be disposed of before that date, University IT Services has a policy of securely wiping network storage infrastructure arrays onsite prior to disposal. Details of the procedure for exporting data from the TRE are covered by document SOP-06-02 Importing and Exporting TRE Datasets. In accordance with the data retention requirements of the Data Controller, the original source data within the TRE can then be destroyed according to document SOP-06-15 Deletion of TRE Datasets. |

Send completed form to datarequest@srft.nhs.uk
How will this be confirmed/ratified

All activity is logged in the Information Management System of Health e-Research Centre and can be confirmed in communications from the Information Security Manager, Ben Green ben.green@manchester.ac.uk. Information Security policies and procedures for the Trusted Research Environment are available on request.

SRFT Comments

Approved by:

| Phil Bell | Jym Bates | Emma Birchall | Signature |

| Date confirmed: |
| Proposed date for data release: |
| Confirmed received date: |

Email sign off

| Appended dates of email |
| Description |

Send completed form to datarequest@srft.nhs.uk
Where will data be stored?

<table>
<thead>
<tr>
<th>Server Name:</th>
<th>Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWEH contact for data release:</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

Send completed form to datarequest@srft.nhs.uk
Variables the six tables of the extracted and linked dataset (provided by Stuart Bennett 13/6/17)

**SIR variables:**

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Journal Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient ID (Pseudo ID)</td>
<td>Patient ID (Pseudo ID)</td>
</tr>
<tr>
<td>LLSOA</td>
<td>ReadCode</td>
</tr>
<tr>
<td>Year of Birth</td>
<td>Rubric</td>
</tr>
<tr>
<td>Sex</td>
<td>EntryDate</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>CodeValue</td>
</tr>
<tr>
<td>Month/Year of Death</td>
<td>CodeUnits</td>
</tr>
<tr>
<td>Source</td>
<td></td>
</tr>
</tbody>
</table>

**EPR variables:**

<table>
<thead>
<tr>
<th>Health Issues</th>
<th>Medications</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient ID (Pseudo ID)</td>
<td>Patient ID (Pseudo ID)</td>
<td>Document Name</td>
</tr>
<tr>
<td>CreatedWhen</td>
<td>Significant Dtm</td>
<td>Display Name (This would be Height (cm) or Weight (kg) for eg)</td>
</tr>
<tr>
<td>ShortName (Health Issue Name)</td>
<td>Medication Name</td>
<td>Value Num</td>
</tr>
<tr>
<td>Onset Date</td>
<td>Frequency Code</td>
<td></td>
</tr>
<tr>
<td>TypeCode (Health Issue Type)</td>
<td>Summary Line</td>
<td>Authored Dtm</td>
</tr>
<tr>
<td>SRFT Extract Date</td>
<td></td>
<td>SRFT Extract Date</td>
</tr>
</tbody>
</table>

Send completed form to datarequest@srft.nhs.uk
SSNAP

- Note that dates in the pathway will be given as number of days relative to onset date (following meeting of Matt Sperrin, Emily Griffiths, and Stuart Bennett 23/6/17).

Patient ID (Pseudo ID)
S1Diagnosis
S1OnsetInHospital
S1OnsetDateTime
S1OnsetTimeNotEntered
S1OnsetDateType
S1OnsetTimeType
S1ArriveByAmbulance
S1AmbulanceTrust
S1CadNumber
S1CadNumberNK
S1FirstArrivalDateTime
S1FirstArrivalTimeNotEntered
S1FirstWard
S1FirstStrokeUnitArrivalDateTime
S1FirstStrokeUnitArrivalTimeNotEntered
S1FirstStrokeUnitArrivalNA
S2CoMCongestiveHeartFailure
S2CoMHypertension
S2CoMAtrialFibrilation
S2CoMDiabetes
S2CoMStrokeTIA
S2CoMAFAntiplatelet
S2CoMAFAnticoagulent
S2RankinBeforeStroke
S2NihssArrival

Send completed form to datarequest@srft.nhs.uk
S2NihssArrivalLoc
S2NihssArrivalLocQuestions
S2NihssArrivalLocCommands
S2NihssArrivalBestGaze
S2NihssArrivalVisual
S2NihssArrivalFacialPalsy
S2NihssArrivalMotorArmLeft
S2NihssArrivalMotorArmRight
S2NihssArrivalMotorLegLeft
S2NihssArrivalMotorLegRight
S2NihssArrivalLimbAtaxia
S2NihssArrivalSensory
S2NihssArrivalBestLanguage
S2NihssArrivalDysarthria
S2NihssArrivalExtinctionInattention
S2BrainImagingDateTime
S2BrainImagingTimeNotEntered
S2BrainImagingNotPerformed
S2StrokeType
S2Thrombolysis
S2ThrombolysisNoReason
S2ThrombolysisNoButHaemorrhagic
S2ThrombolysisNoButTimeWindow
S2ThrombolysisNoButComorbidity
S2ThrombolysisNoButMedication
S2ThrombolysisNoButRefusal
S2ThrombolysisNoButAge
S2ThrombolysisNoButImproving
asS2ThrombolysisNoButTooMildSevere
S2ThrombolysisNoButTimeUnknownWakeUp
S2ThrombolysisNoButOtherMedical
S2ThrombolysisDateTime
S2ThrombolysisTimeNotEntered

Send completed form to datarequest@srft.nhs.uk
Send completed form to datarequest@srft.nhs.uk
Send completed form to datarequest@srft.nhs.uk
Send completed form to datarequest@srft.nhs.uk
S7DischargePIConsent
S8FollowUp
S8FollowUpDate
S8FollowUpType
S8FollowUpBy
S8FollowUpByOther
S8FollowUpPIConsent
S8MoodBehaviourCognitiveScreened
S8MoodBehaviourCognitiveSupportNeeded
S8MoodBehaviourCognitivePsychologicalSupport
S8Living
S8LivingOther
S8Rankin6Month
S8Rankin6MonthNK
S8PersistentAtrialFibrillation
S8TakingAntiplateletDrug
S8TakingAnticoagulent
S8TakingLipidLowering
S8TakingAntihypertensive
S8SinceStrokeAnotherStroke
S8SinceStrokeMyocardialInfarction
S8SinceStrokeOtherHospitalisationIllness

Send completed form to datarequest@srft.nhs.uk