

How routine Accident and Emergency (A+E) attendance data improves ambulance service resilience

Data Collection

What data will you collect or create?

13 hospital trusts will input the following data in a Python based web form on a monthly basis:

- period of data collection - date
- unique organisation code - string
- number of attendances during period - integer
- number of four hour breaches during period - integer
- number of attendances resulting in hospital admission - integer
- consent to process data - boolean

The data collection tool will then save data into a secure .csv file for later analysis in RStudio.

How will the data be collected or created?

A purpose built data collection tool has been developed making use of Jupyter widgets in a Python based web interface. The data collection tool ensures only the necessary information is uploaded and will only save records for which consent has been confirmed by the user. The tool is designed in a manner that no patient identifiers can be captured. Data will be saved on completion of the tool in .csv format and resultant files will be version controlled monthly using Github in a well developed folder structure.

Data entry will only be available to permitted individuals that have been given appropriate training and guidance in its use.

Documentation and Metadata

What documentation and metadata will accompany the data?

Metadata will include a data dictionary - produced in RStudio.

All version Git pushes will state clear information regarding what changes have been made and why.

Ethics and Legal Compliance

How will you manage any ethical issues?

A consent checkbox will require completing by those performing data entry via the data capture tool. This will ensure consent has been received to share data.

The only data to be collected are aggregated hospital trust attendance figures. No individual patient identifiers will be passed into the data capture tool.

Data will be stored for as long as is required by any ethics/funding bodies and the Data Protection act 2018 before safe destruction.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Question not answered.

Storage and Backup

How will the data be stored and backed up during the research?

The raw dataset is relatively small, with 78 data points being generated each month into a .csv file of 13 rows and 6 columns. Data will be stored securely in an approved cloud based server with automatic backups saved locally in a secure environment.

How will you manage access and security?

Access to data will only be granted to approved individuals that require it for the proposed use and for monitoring purposes. Approved users will be required to use login credentials that have been granted access, making use of two-factor authentication. Access to the database will be timestamped with user details.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

As the data are fairly generalisable and non-identifiable, data will be stored online and made open source in .csv format. The organisation codes will be replaced with a unique identifier (eg. trust_1, trust_2 etc.) to keep them anonymous. All saved records will have had the consent box checked so all that will be stored is period, unique_trust, attendances, breaches and admissions data - the performance columns can easily be calculated at a later date if whoever is accessing the data wishes but is unnecessary to store at this point.

What is the long-term preservation plan for the dataset?

The dataset will be archived at the end of study in accordance with university, funder, legal and ethical regulations

Data Sharing

How will you share the data?

Data will be shared by means of end of study reports, posters, conferences and peer reviewed journals.

Requests specifically for access to the dataset will be considered on a case by case basis once the initial study period is over and final report written. Sharing of any datasets will involve secure methods to ensure data are passed safely.

Are any restrictions on data sharing required?

Hospital trusts may well require a data sharing agreement to be drawn up before inputting any data to the data capture tool. This should not cause too much of a problem as there are minimal data points, no staff or patient identifiers and the data to be shared is aggregated data that is already reported elsewhere. However, it is well appreciated that these processes can take considerably longer than expected in NHS organisations due to limited staffing and strict cyber-security procedures.

Responsibilities and Resources

Who will be responsible for data management?

The PI will be responsible for all aspects of data management as the study is not particularly complex and time consuming and there is a relatively small amount of data to process. If it is deemed necessary, roles may be delegated to a trained individual at a later date, with appropriate

documentation to indicate who has been delegated a particular role (eg. delegation log).

What resources will you require to deliver your plan?

The study will run from an existing laptop (unless roles delegated as previously discussed). All code will be written using opensource Jupyter notebooks running Rstudio and Python. An existing Microsoft 365 license will be used to access SharePoint for internal file management and .csv files may also be opened using Microsoft Excel throughout the study.