The University Of Sheffield.

NHS Yorkshire Ambulance Service NHS Trust

Safety INdEx of Prehospital On Scene Triage (SINEPOST)

The derivation and validation of a risk prediction model to support ambulance clinical transport decisions on scene.

Study Protocol

This protocol has regard for the HRA guidance and order of content

Research Reference Numbers

Protocol version number and date

Protocol version 1.1 04/12/2019

IRAS Number:	260505
SPONSORS Number:	YASRD109
FUNDERS Number:	ICA-CDRF-2018-04-ST2-044

Signature Page

The undersigned confirm that the following protocol has been agreed and accepted and that the Chief Investigator agrees to conduct the study in compliance with the approved protocol and will adhere to the principles outlined in the Medicines for Human Use (Clinical Study's) Regulations 2004 (SI 2004/1031), amended regulations (SI 2006/1928) and any subsequent amendments of the clinical study regulations, GCP guidelines, the Sponsor's (and any other relevant) SOPs, and other regulatory requirements as amended.

I agree to ensure that the confidential information contained in this document will not be used for any other purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of the Sponsor

I also confirm that I will make the findings of the study publically available through publication or other dissemination tools without any unnecessary delay and that an honest accurate and transparent account of the study will be given; and that any discrepancies and serious breaches of GCP from the study as planned in this protocol will be explained.

For and on behalf of the Study Sponsor:

Signature: Fiona Bell

Date: 4/12/19

Name (please print): Dr. Fiona Bell

Position: Head of Research and Development

Chief Investigator:

Signature:

Date: 5/12/19

Name: (please print): Professor Suzanne Mason

Key Study contacts

Chief Investigator (and primary supervisor)	Prof. Suzanne Mason University of Sheffield CURE, ScHARR 30 Regent Street S1 4DA Email: S.Mason@Sheffield.ac.uk
Student Investigator	Jamie Miles Yorkshire Ambulance Service NHS Trust Springhill, Brindley Way Wakefield 41 Business Park Wakefield WF2 0XQ
Sponsor	Yorkshire Ambulance Service NHS Trust Springhill, Brindley Way Wakefield 41 Business Park Wakefield WF2 0XQ
Funder(s)	National Institute of Health Research Academy 21 Queen Street Leeds LS1 2TW Tel: 0113 532 8444 Email: academy-awards@nihr.ac.uk
Statistician	Dr. Richard Jacques Medical Statistics Group School of Health and Related Research (ScHARR) The University of Sheffield Regent Court, 30 Regent Street SHEFFIELD, S1 4DA Email: R.Jacques@sheffield.ac.uk

i. List of Contents

ii. List of Abbreviations

BI	Business Intelligence (Yorkshire Ambulance Service)
CAG	Confidentiality Advisory Group
CI	Chief Investigator
CiCS	Corporate Information and Computing Services (University of Sheffield)
CURE	Centre for Urgent and Emergency Care Research (University of Sheffield)
DOLS	Deprivation Of Liberty Safeguards
DSA	Data Sharing Agreement
ECDS	Emergency Care Data Set (NHS Digital)
ED	Emergency Department
ePCR	Electronic Patient Care Record (Yorkshire Ambulance Service)
EPV	Events Per Variable (aka Events Per Parameter)
FAST	Face, Arms, Speech, Time
GCP	Good Clinical Practice
HRA	Health Research Authority
ICMJE	International Committee of Medical Journal Editors
IMD	Index of Multiple Deprivation
LASSO	Least Absolute Shrinkage and Selection Operator
LOESS	Locally weighted Scatterplot Smoothing
LSOA	Lower Super Output Area
NACN	National Ambulance Commissioners Network
NEWS	National Early Warning Score
NHSD	NHS Digital
NIHR	National Institute of Health Research
NPV	Negative Predictive Value
PPV	Positive Predictive Value
RDS	Research Design Service
ROC	Receiver Operating Characteristic
ScHARR	School of Health and Related Research (University of Sheffield)
SECF	Sheffield Emergency Care Forum
SHARRP	Sheffield Addiction Recovery Research Panel
SINEPOST	Safety Index of Prehospital On Scene Triage
SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
UECRT	Urgent and Emergency Care Review Team
UOS	University of Sheffield
VM	Virtual Machine
YAS	Yorkshire Ambulance Service

iii. Study Summary

Study Title	The Safety Index of Prehospital On Scene Triage: The derivation and validation of a risk prediction model to support ambulance clinical transport decisions on scene.
Internal ref. no. (or short title)	The SINEPOST study.
Study Design	Risk prediction modelling on linked healthcare data.
Study Participants	Adult (over the age of 18) ambulance service patients attended by Yorkshire Ambulance Service between July 2019 and December 2019.
Planned Sample Size	466,846
Planned Study Period	September 2019 – June 2021

	Objectives	**Outcome Measures**
Primary		
Can ambulance service clinical data predict an avoidable attendance at the ED in adults using classification models?	To build classification models deriving risk predictions using prehospital clinical data as input variables, and ED experience as the output variable.	ED experience (avoidable/non-avoidable)
Secondary		
What is the simulated transportability of the model derived from the primary outcome?	Internally validate the model and apply to a retrospective cohort of non-conveyed patients.	ED experience (avoidable/non-avoidable)
Which classification model is most accurate at predicting an avoidable attendance at the ED in adults?	Compare the different classification models for most accurate and feasible to embed in practice.	Comparative accuracy in prediction, misclassification rate, imprecision and inaccuracy.

iv. Funding and support in kind

Funder(s)	Financial and non-financial support given
National Institute of Health Research Academy 21 Queen Street Leeds LS1 2TW Tel: 0113 532 8444 Email: academy-awards@nihr.ac.uk	£218,719

v. Role of study sponsor and funder

Yorkshire Ambulance Service NHS Trust (YAS) is the sponsor of this study as the employer of the student investigator. Yorkshire Ambulance Service will undertake all sponsor responsibilities outlined the UK Policy Framework for Health and Social Care Research.

The National Institute of Health Research (NIHR) Academy are the funder of this study. They will undertake all funder responsibilities outlined in the UK Policy Framework for Health and Social Care Research.

vi. Roles and responsibilities of study management groups & individuals

There will be no Study Management Committees

vii. Protocol contributors

This protocol has been primarily drafted by the Student Investigator Jamie Miles.

There has been academic input from Professor Suzanne Mason, Dr Richard Jacques, Janette Turner and Maxine Kuczawski who helped develop the methodology and data management sections of the protocol

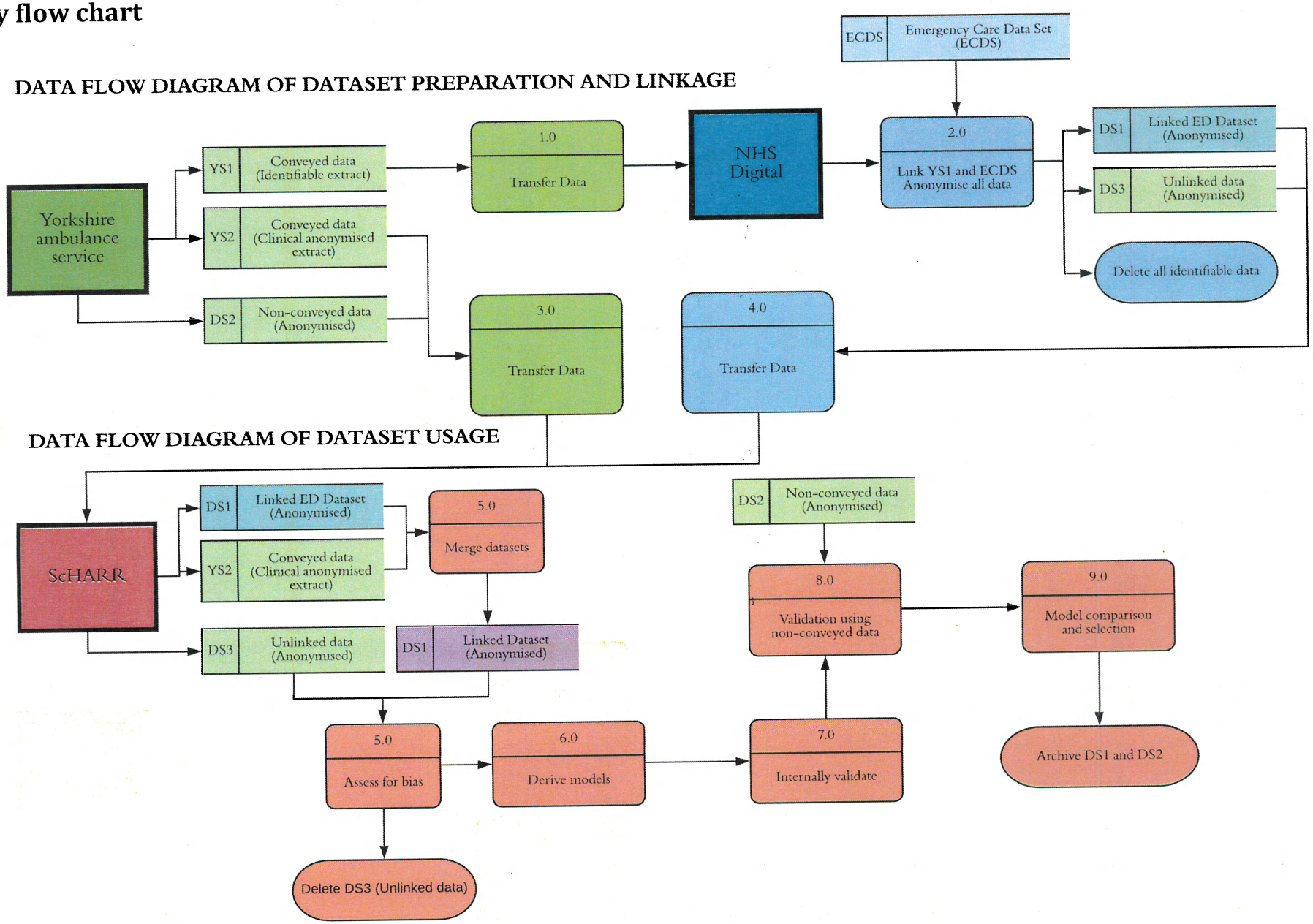
There has been sponsor input from Jane Shewan and Dr Fiona Bell who helped develop the study management and design sections of the protocol.

There has been a Public Involvement forum set up for this project. They have helped develop the lay sections of the protocol.

viii. Key words

Risk Prediction, Ambulance Service, Discharge on scene, Data linkage, Modelling

ix. Study flow chart



1 Background

The demand problem

Demand for the urgent and emergency care system is increasing (~5.2% annually).[1] 10.7 million calls were transferred to the ambulance service with 6.6 million receiving a face-to-face attendance.[1] The vast majority of these calls were urgent and not emergencies. This demand is also mirrored in ED, which results in delays when the ambulance service is handing the patient over. In a 3 month winter period in 2017/18 in England, there were 41,879 patient episodes where an ambulance took over an hour to handover to the ED.[2] There is international evidence from the USA[3-6], Canada[7], UK[8] and Australia[9-11] showing ambulance ramping can be both a consequence and a cause of ED crowding. Queuing ambulances means less resources are available to respond to 999 calls, which has a damaging effect on patient care and ambulance service performance.[12] Furthermore, the accuracy in the decision to convey a patient to the ED is not always accurate. A 2018 study found that up to 16.9% of ambulance conveyances to the ED were potentially avoidable.[13] This study used a process-based validated definition of avoidable, which is being used in this project as the primary outcome measure.[14,15]

More evidence is required in identifying patients who are likely to have an avoidable conveyance to the ED, whilst they are still with the ambulance service on-scene.

The case-mix problem

The majority of ambulance service patients require less critical interventions and more community based care.[16] In response to this, policy has been driving paramedics to make autonomous and dynamic decisions about their patient care before transporting them to the ED.[16-19] This theme was particularly emphasised in the Urgent and Emergency Care Review Team (UECRT) report in 2013.[17] One of the main outcomes they aimed to achieve in the report were for those with an urgent need to access healthcare be provided with a highly responsive, effective, and personalised service outside of hospital.

Interventions have been developed to stratify patient risk on-scene and decide where the best care of a patient should be. One example that has been adopted by numerous ambulance services is the paramedic pathfinder tool.[20] The tool was user tested in 2014 when a sample of 481 patients had the tool applied to them with results showing sensitivity of 94.83% and specificity of 57.9%.[21] The tool was highly risk averse with almost half of the patients being transported to the ED, when the care could have safely been elsewhere. Furthermore, O Hara et al. used qualitative methods to conclude that paramedic decision making is multifactorial and complicated.[22] They emphasised a decision not to transport a patient was the most challenging.

More evidence is required to risk stratify patients and allow paramedics to make more discriminate ED transport decisions on-scene.

2 Rationale

This research is supporting paramedics to make more appropriate and effective decisions for patients who may not require the level of care provided by a hospital. It is important as it is aiming to navigate care decisions that will safely provide patients with the right care, first time. If a paramedic can see the likelihood that their patient may have an avoidable attendance, it opens up an opportunity to explore community options. It also empowers the patient to be an active partner in developing a self-care plan. The whole urgent and emergency care system would benefit from this research. If paramedics can make safe non-transport decisions it frees their capacity to assess a patient still waiting for help. With less patients being transported to the ED with low acuity problems, it will minimise delays in caring for those who do need specialist ED interventions. For community care it will identify patients that require out of hospital services so that they can get the right level of care promptly. This project does recognise when a paramedic assesses a patient, there are six possible outcomes relating to a transport decision. Firmer evidence surrounding outcome measures for the other five dispositions is needed before tools that can accurately signpost patients to any service can be developed.

Review of existing evidence

Paramedic decision making

In 2014, O'Hara et al. published a qualitative report on paramedic decision making in transitioning care.[22] It appears to be the most comprehensive exploration on the subject to date. They identified the most complex type of decision surrounds non-conveyance and the concept of discharging the patient on-scene. This conclusion agrees with many other sources.[23-26] A cause of this complexity is the perception of job role as illustrated by Hoikka et al. They elude to paramedic education being focussed on high acuity situations, which foster a culture that could struggle in recognising and discharging low acuity patients.[27] Simpson et al. also demonstrated that role perception was crucial in the decision-making process of paramedic's on-scene. They undertook a qualitative study with thirty-three paramedics examining decision-making in the elderly who have fallen. They acknowledged that role perception was profound on how a paramedic approaches a decision.[26] This idea is reaffirmed by Brydges et al. who state:

"Many of the participants perceived their role as a paramedic to be defined by responding to emergency calls for help (i.e., consistent with their initial education and certification expectations), and referral programs represented a formal departure from that enduring view." (p. 633)[28]

Non-conveyance

Non-conveyance can be defined as "an ambulance deployment as appropriate, where the patient after examination and/or treatment on-scene does not require conveyance with medical personnel and equipment to the healthcare facility".[29] Ebben et al. produced a recent systematic review on the subject; focusing on patient safety and outcomes further along in the patient journey.[30] This systematic review appears to be the most comprehensive in comparison to other reviews undertaken in the same subject area.[31-34] Sixty seven studies were included in the review with the majority being quantitative observational studies. The rate of non- conveyance varied between 3.7%–93.7%. This was due to the differences in prehospital models between countries.

According to NHS England, the current non-conveyance rate experienced nationally in the UK is 37.6% (locally in Yorkshire it is 30.6%).[35] The limitation of this report is that the data collected is used as a performance measure for the ambulance service. This could introduce a bias into the data collection stage and mar the accuracy of the report.[36] A different approach to non-conveyance would be to ascertain clinical necessity of conveyance. Patton and Thakore found that a third of their sample attended the ED 'inappropriately' although their sample size was low (n=295).[37]

Decision tools

There is currently a lack of evidence to support the use of tools to assist in prehospital decision making regarding the need to transport the patient to the ED. Conversely, tools used within the ED to risk stratify patients and predict admission have demonstrated high accuracy.[38-41] Cameron et al. compared two admission prediction tools: the Glasgow Admission Prediction Score (GAPS) and Amb Score.[42] They found the Area Under Curve (AUC) of 0.807 (95% CI 0.785-0.830) for GAPS and 0.743 (95% CI 0.717 to 0.769) for the Amb Score. These were derived using multivariable logistic regression. Advances in technology, in particular electronic health records has enabled complicated predictive models to support the clinician in making more discriminate decisions.[43-46] Levine et al. used random forest modelling to triage patients entering the ED and compared this with the Emergency Severity Index (ESI).[46] They found the random forest was more accurate at stratifying mid-severity patients with an AUC ranging between 0.73-0.92 across all levels of the ESI.[46]

3 Objectives and Outcome Measures/Endpoints

3.1 Primary research question

Can ambulance service clinical data predict an avoidable attendance at the ED in adults using classification models?

3.1.1 Primary objective

To build classification models deriving risk predictions using prehospital clinical data as input variables, and ED experience as the output variable.

3.1.2 Primary outcome measure

An avoidable attendance at ED as defined by O'Keeffe et al. (2018).[15]

3.2 Secondary research questions

What is the simulated transportability of the model derived from the primary outcome?

Which classification model is most accurate at predicting an avoidable attendance at the ED in adults?

3.2.1 Secondary objectives

Internally validate the model and apply to a retrospective cohort of non-conveyed patients.

Compare the different classification models for most accurate and feasible to embed in practice.

3.2.2 Secondary outcome measure

Avoidable attendance at ED as defined by O'Keeffe et al. (2018)[15]

Comparative accuracy in prediction, misclassification rate, imprecision and inaccuracy.

4 Study Design

This prognostic study has a retrospective correlational design. Phase 1 will start by receiving a dataset of ambulance service journeys linked to ED records. This will be linked and anonymised externally by NHS digital (NHSD). Each patient episode in the dataset will contain clinical information from both the ambulance service and ED. This linked cohort will be known as Cohort 1. Any unlinked data will be returned for an assessment of selection bias.

Classification models will be applied to the dataset to predict an avoidable attendance at ED (primary outcome measure). Phase 1 will end when the models have been derived.

Phase 2 will use methods such as bootstrapping and cross-validation to internally validate the models. The models will also be applied to random samples of patients who were not conveyed to the ED. Phase 2 will end when the most accurate model has been identified and selected.

5 Study Setting

This study is set within the boundaries of Yorkshire Ambulance Service. YAS serves a population of over five million people and covers 6000 miles of varied terrain from the isolated Yorkshire Dales and North York Moors to urban areas including Bradford, Hull, Leeds, Sheffield, Wakefield and York.

6 Participant Eligibility Criteria

6.1 Inclusion criteria

Inclusion 1 (for Cohort 1 ED information)	• Age 18 years old or older. • Transported to ED by Yorkshire Ambulance Service between July 1st 2019 and December 31st 2019. • Have an ED Care record of the event.
Inclusion 2 (for Cohort 1 Ambulance information)	• Age 18 years or older. • Assessed by a qualified ambulance clinician ((either paramedic (of any level) or technician grade II)). • Had an electronic patient care record completed. • Transported to an ED between July 1st 2019 and December 31st 2019. • Were handed over and booked in as a patient to the ED
Inclusion 3 (for Cohort2)	• Age 18 years or older. • Assessed by a qualified ambulance clinician (either paramedic or technician grade II). • Had an electronic patient care record completed. • Discharged on scene and not transported between July 1st 2019 and December 31st 2019.

6.2 Exclusion criteria

Exclusion 1 (for ED cohort)	• Patient cases where they were less than 18 years old at time of episode. • Patient cases where there were five or more attendances within the data collection period.
Exclusion 2 (for Ambulance Service cohort 1)	• Patient cases where they were less than 18 years old at time of episode. • Patient cases where they had five or more patient contacts within the data collection period.
Exclusion 3 (for Ambulance Service cohort 2)	• Patient cases where they were less than 18 years old at time of episode. • Patient cases where they had five or more patient contacts within the data collection period. • Patient cases that were transported by the ambulance crew on scene.

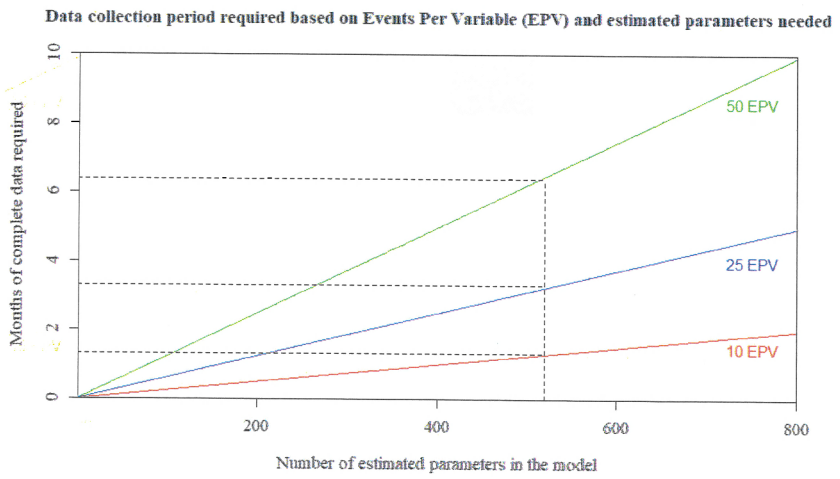
7 Statistics and Data Analysis

7.1 Data linkage plan

All patients (using the cohort 1 eligibility criteria) will be retrieved from the Yorkshire Ambulance Service data warehouse by the Business Intelligence (BI) team within the organisation. This data will not be seen by the CI or student investigator. Once retrieved, the data will be split into two different extracts. The first is the clinical information extract. This will be anonymised completely and delivered securely to the School of Health and Related research (ScHARR) at the University of Sheffield. The second is the identifiable information extract. This will be securely sent to NHSD. They will link this data in a secure environment with the associated ED record for each patient. The methodology used is transparent and well documented.[47] Once maximum data linkage has been achieved, the dataset will be completely anonymised and divided into 'linked' and 'unlinked' cases. This is because it is anticipated that not all cases will be able to be linked. These two datasets will then be securely delivered to the ScHARR and stored appropriately in accordance with the University of Sheffield (UOS) Information Governance policies. All identifiable information will remain in NHSD and will be deleted as soon as the data is linked. The linked data will then be merged with the clinical extract. The unlinked data will be used for a bias assessment at UOS before being completely deleted. The data flow diagram for this project can be found in section ix above.

7.2 Sample size calculation

Numerous studies have suggested that each candidate predictor requires at least 10 Events Per Variable (EPV).[48-54] All candidate predictors (twenty three) have been transposed from the YAS electronic Patient Care Record (ePCR) and can be found in table 1. These output a liberal estimation of 520 parameters, and therefore 5200 events. A study examining avoidable conveyances reported a lower limit of 9% avoidable conveyances in the same population as this study.[13] When this is applied to the minimum number of events, it would yield a sample size of 57,778. However, one study has shown that up to 50 events per variable are needed for variable selection.[55] This would alter the sample size to 288,889. Data linkage on the same data had a success of 80% using deterministic matching on hierarchical variables. There was also 7.7% of data which was incomplete to create the outcome variable (avoidable attendance).[13] This would alter the potential sample size for the first cohort to 368,911. In 2018, YAS assessed 919,635 patients on scene and conveyed 735,053 (74.9%). This would mean a data collection period of 6 months in order to fulfil the sample size of 368,911.



7.3 Model Development

7.3.1 Data preparation and missing data for cohort 1

There will be an initial exploration to examine quantity of candidate variables within the dataset and the parameters within each one. This exploration will also identify missing data, and highlight any variables that may need cleaning or preparing in anticipation of model development. The outcome variable will be created using the variables derived from the ED data. Once created, the ED variables will be removed from the dataset. This completes the data preparation stage.

Any missing data will be assessed for randomness using statistical methods such as Little's Missing Completely at Random Test. If there is a small (<2%) of random data missing, then complete case deletion of the missing data will be considered. Alternatively, any arbitrary missing data will be handled with multiple imputations. Any monotonic pattern of missing data will be approached after further consultation with a statistician to prevent systematic bias entering into the models.

Following on from this, the cohort 1 sample will be compared to the unlinked data to assess for screening bias. Once this is complete, the unlinked data will be destroyed.

Predictive models will be developed and applied to the data independently. There will be a preference to use the whole dataset to derive each model and then use resampling techniques such as bootstrapping to measure optimism and recalibrate, (as opposed to sample splitting). For models which do require a training set and test set, this split will be done non-randomly, with the test data being unlabelled.

7.3.2 Predictors

There will be no a priori candidate variables selected. Instead, all clinical variables collected during the paramedic patient assessment and/or treatment will be included in the initial development. Lower Super Output Area (LSOA) will be recoded into an Index of Multiple Deprivation (IMD) score which will be used as a predictor. It is anticipated there will be twenty three variables across three domains as described by table 1. Penalisation will be used in logistic regression to eliminate variables which have little or no effect on the outcome. Such methods include Ridged, LASSO and Elastic Net.

Table 1: Example of possible candidate predictors

Physical factors	Social factors	Interventional factors
Age	Location	Airway
cABC (initial assessment)	Dementia	Drug administration
Working impression	Communication	Cannulation
Allergies	DOLS	Resuscitation
Medication	Mental capacity	Extrication
Primary NEWS	Moving and handling aids	Clinical procedures
Subsequent NEWS	Mobility	Patients advice
FAST	IMD	

7.3.3 Data Analysis

For each model, a contingency table will be devised so summary accuracy statistics including sensitivity, specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) can be calculated. In addition, a calibration plot will show the agreement between observed and expected results, with a LOESS smoother and confidence intervals being used. Discrimination will be

calculated with the C-statistic, plotting on a Receiver Operating Characteristic (ROC) curve, or by further examination of the calibration plot. These results will all be unadjusted at the derivation stage and the analysis will only show the apparent performance of each model.

7.4 Secondary outcome analysis

7.4.1 Sample size for cohort 2

This will be the maximum number of non-conveyed patients that can be attained in the same data collection period as cohort 1. A YAS report has established there were 93,271 non-conveyed patients between July and December 2018. With a 5% demand increase for 1 year forecasting, and assuming there will be no change in the percentage of non-conveyed, the sample size estimation would be 97,935.

7.4.2 Data preparation and missing data for cohort 2

This is the unlinked ambulance service data and only includes patients that were not conveyed to hospital. Its purpose is to sense check the derived models performance in a sample of low-mid acuity patients. All variables will be explored with a frequency analyses undertaken to identify missing data. The same preparation and methods for dealing with missing data will take place in the ambulance clinical variables as cohort 1 (see protocol section 10.3.1). No outcome will be generated as it is unlinked.

7.4.3 Internal validation

For the logistic regression model this can be performed using bootstrapping of cohort 1 to measure the optimism and then apply the shrinkage factor to the beta coefficients. For methods such as random forest modelling, they would need alternative methods such as cross-validation. Once all models have been internally validated, they will be applied to cohort 2.

There will be a sub-group analysis using non-random splitting of cohort 1 according to individual variables. The performance in the sub-groups can then be plotted in a graphical way such as a calibration plot or ROC curve. This will enable the identification of groups of patients the model is more vulnerable/ robust towards.

7.4.4 Model selection

Once all adjusted models have been created and tested on cohort 2, there will be a comparative analysis of model performance using the same methods described above. The CI will then present the models to a small expert panel for independent confirmation. This panel will be a combination of clinicians and statisticians.

8 Data Management

8.1 Data collection tools and source document identification

Yorkshire Ambulance Service has an ePCR system in place across the whole region. This is a compulsory system for paramedics to create and complete individual care records for each patient which begins at arrival on scene, and is completed once the patient is handed over. This document forms the source document for each patient that will be eligible in this project. The sample size in section 10.2 is in effect calculating the number of source documentation required for model derivation and validation.

Within each source document is the variables required to derive and validate the model. An example of the variables within can be seen in table 1 (above). The variables are mainly categorical and presented as drop down lists. The other type of variable captured is continuous such as age. The clinician is expected to fully complete each record however not all fields are compulsory capture.

This is a retrospective study and as such, there will be no attempt to maximise completeness of data. The sample size has been inflated to account for expected source document attrition at the point of data retrieval. The methodology has accounted for the handling of missing data.

8.2 Data handling and record keeping

8.2.1 Data linkage

This study is using routine ambulance service data and linking each episode to the matched ED record to create the final dataset. The initial data retrieval will be done by the data controllers of the ambulance service data. A final data specification will be submitted to ensure only the required data is retrieved. On retrieval, a unique ID will be created for each ePCR and entered as a new variable. Once this has happened, the data will be split into a clinical extract (which is anonymised) and an identifiable extract. The unique ID will be present in both extracts.

The clinical extract will be transferred securely to the CI, and the identifiable extract will be transferred to NHSD for data linkage. A Data Sharing Agreement (DSA) will be created between NHSD and Yorkshire Ambulance Service for the transfer of data.

NHSD will then link the ambulance service record to the associated ED record. This will be done using hierarchical matching of key variables. Once maximum data linkage has taken place, NHSD will delete all identifiable data from the dataset, leaving only the unique ID. This anonymised dataset will be transferred securely to the CI. Once the CI has both datasets, they can merge them using the unique ID. Please see the study flow chart above.

8.2.2 Data management

Once the data has been linked, anonymised, and sent securely to the CI, the data will be stored securely at the UOS. This will be in concordance with the Information Governance policies and procedures.

The storage of the dataset will be held and accessed only by the CI and their supervisory team, ensuring confidentiality and integrity of the data. All computers used to host servers are kept in secure premises where access is restricted to authorised individuals. No data is or will be stored on laptop computers. There is a strict policy of not accessing data on the network from laptop computers whilst in insecure, public locations.

The dataset will be stored on a Virtual Machine (VM) which is secured and maintained by The UOS Corporate Information and Computing Services (CiCS). CiCS administrators have policies in place which address network security (especially threats from outside the campus network) and software maintenance. The VM is privately addressed; it is not accessible from outside the campus network. Access to the VM (only possible from within the campus network and from specified IP addresses) is granted only to a limited number of user accounts, all of which require authentication by username and password. The machine-level firewall policies permit incoming traffic only from specified IP addresses.

8.2.3 Access to data

Only the CI and the direct supervisory team will have access to the data within the UOS.

8.2.4 Archiving

Storage of the dataset and associated folders are in folders designated with a 'time to live', projected to be a period of 10 years. After the 10 year period has expired, the research database will be removed from the study-specific system folder and archived in a secure UOS networked storage area, accessible only to administrators of the storage network.

9 Monitoring, Audit & Inspection

This study may be audited as part of the routine audit process as laid out in the sponsors research governance policy. All source documents and essential documents will be available to the sponsor for audit for at least 10 years.

10 Ethical And Regulatory Considerations

10.1 Research Ethics Committee (REC) review& reports

Ethical approval will be sought from the Health Research Authority (HRA) including support from the Confidentiality Advisory Group (CAG). This is for obtaining identifiable patient data without consent. All amendments will be submitted to the HRA for approval prior to implementation. All correspondence with the HRA will be recorded and stored in the local study folder, which can be accessed by the sponsors. An annual progress report will be submitted to the HRA within 30 days of the anniversary date on which the favourable opinion was given, and annually until the study is declared ended. It is the CI's responsibility to produce the annual report and also to notify the HRA of the end of the study.

10.2 Peer review

This study has had a high quality peer review by the funder, as well as the sponsor.

10.3 Public and Patient Involvement

This project has been heavily influenced by the public. The Sheffield Emergency Care Forum is a Public Involvement and Engagement forum that is situated at the UOS. For the 12 months leading up to funding submission, this project has had a standing agenda at their quarterly meetings. Through an iterative process, they have steered the design of the project.

A small grant from the Research Design Service (RDS) was won to conduct three Public Involvement events. There were a total of twenty two public members involved who captured diversity in socio-economic status, gender, age, disability and ethnicity.

The members of the public who helped in the design of the research have been invited to a Whatsapp group. This is a digital messaging application that is secure with end-to-end encryption. Those who have access to the application but never used it will be helped. Those who don't have access but would like to participate will be invited to face to face meetings and caught up on a one-to-one basis. The purpose of the Whatsapp group is to have a longitudinal conversation with the public about the project. It is understood that the PhD is a learning experience and highly likely to change as the project progresses. As these changes are being made, the public will be asked in advance of their thoughts. They will also be invited to comment on reports to ensure the language can be understood by everyone. There will be annual meetings with members of the group to consolidate conversations in the Whatsapp group, and also to present and discuss the progress of the project. In the first phase of the research the panel will be invited to help produce a lay summary of the risk prediction model. This phase could be considered esoteric and complicated and the public will be invaluable in ensuring the summary is clear and concise.

In the second phase, there will be an opportunity to present progress with the study, overview the tool and how it works. This will also provide an opportunity for discussion and views to be shared about the tool. The output from this consultation will be used to influence the ongoing development and implementation of the tool. Willing volunteers will be asked to create a video for the public introducing the tool and its purpose. They will also be invited to co-author relevant publications, and abstracts will be submitted to INVOLVE and other conferences, offering them an opportunity to present.

The public's acceptability of using patient identifiable data in this study without consent has been tested on three separate occasions with three different public involvement groups. The SEFC held a discussion about the project on the 10/08/2018 and there was a consensus of support from the group about using identifiable data

in this study without consent. On the 13/08/2018 a random group of public members were identified using an NIHR Public Involvement grant. They felt the method of using identifiable data without consent was justified and supported the methodology. On the 16/08/2018 the Sheffield Addiction Recovery Research Panel (SHARRP) were consulted and felt the research question was important and that the methodology was appropriate as large numbers were required to create a model.

10.4 Protocol compliance

Any breach of protocol will be reported via the YAS Datix system to notify the Research and Development department at YAS. Duty of Candour will be considered in consultation with the YAS Duty of Candour specialist. Furthermore, the UOS Information Governance Committee will be informed and will undertake a further investigation of the breach.

10.5 Data protection and patient confidentiality

This protocol complies with each principle of the Data Protection Act 2018. Below is a summary statement for each principle.

10.5.1 Fair Processing

The legal basis for collecting and processing of patient identifiable data is that it is a task in the public interest.

The data is being used for linking and validating data from different healthcare providers in order to derive and validate a risk prediction model to support decision-making.

Sensitive patient data is being processed under the condition that the data is processed for medical purposes by individuals subject to an equivalent duty of confidentiality as exercised by a health professional.

The data processing is being used only for this legitimate reason and the use of data to link healthcare provider records will not have any adverse effects on individuals whose data is linked.

This data is being collecting without the consent of patients, however this is justified on the grounds consistent with Article 4 of the Data Protection (Processing of Sensitive Personal Data) Order 2000, which states that processing data without consent may be justified when in substantial public interest and is necessary in a case where the data controller cannot reasonably be expected to obtain the explicit consent of the data subject.

10.5.2 Used for specified purposes

(Not required)

10.5.3 Minimum necessary for the purpose

There will be the minimum amount of data collected which will satisfy the required sample size for the study. Only variables that will have potential to act as a candidate predictor, or are necessary for data linkage will be requested.

There will be the linkage of data from YAS, and data held by NHSD. Only data items of name, address including postcode, date of birth, age, gender, NHS number and incident/episode number are required for data linkage. A more limited number of fields are not possible to specify or collect as the nature of YAS data is such that for different cases different personal identifiers will be missing i.e. for some YAS cases, no home address will be specified as the incident may have happened away from the individuals' home. In this case age, gender and name/ NHS number will be required to enable linkage.

10.5.4 Accuracy

It will be ensured that all information collected and stored from the service providers is recorded accurately and reflects the source data. There will be a dependence on the accuracy and completeness of data collected by service providers at time of episode. There will only be data retrieval for linkage of specific episodes which have occurred on a specific date and there will not be any follow up or contacting of patients at a subsequent date. Thus there will be no contemporaneous need to ensure records are kept up to date.

10.5.5 Kept for minimum time necessary

As soon as the data is linked, all identifiable data will be destroyed by NHSD.

10.5.6 In accordance with the rights of data subject

This study is requesting permission to access data without consent of patients. The study design has been carefully considered to protect the personal and identifiable information of service users. As such, none of the research team will receive this sensitive information.

There will be no direct marketing undertaken, this data is being processed for the defined process of data linkage to form a large dataset containing de-identified data, which is being used to develop a predictive model.

Regarding holding inaccurate data – it is possible that data will become out of date as the study is undertaking retrospective data analysis. Therefore inevitably some patients will move address or have other changed circumstances. However the research team will ensure that the data recorded and processed is accurate at the time i.e that it accurately reflects the data provided by the healthcare providers.

10.5.7 Security and confidentiality protection

Both YAS and NHSD have robust systems in place for the storage, retrieval and processing of patient identifiable information. In addition, they hold the knowledge and expertise to ensure the security of sensitive data. Identifiable data will only be seen by the BI department in YAS who will separate the context (clinical information) from the identifiable data. This protects patient confidentiality. It also ensures NHSD only process the minimum amount of variables for data linkage. As soon as NHSD have linked the data, the identifiers will be destroyed. This anonymised dataset can then be securely transferred to the UOS.

The CI and their supervisors are employees or part of The UOS and are subject to The UOS Information Security Policies. This includes complying with the Data Protection Act 2018 and UOS Information Governance Policy which requires renewal every 2-years. The CI and the sponsors are bound by the Information Governance policies and procedures of Yorkshire Ambulance Service and are regularly re-certified on these.

Physical security is maintained by University staff working in secure locked offices in alarm coded corridors. Staff members trained to log off from their computers whenever they are away from their desks

10.5.8 Not disclosed outside of the EU

No patient or sensitive data on individuals will be disclosed outside the EU.

10.6 Financial and other competing interests

No competing interests

10.7 Amendments

Any amendments to the protocol will be submitted to the HRA for categorisation and consideration as appropriate.

10.8 Access to the final study dataset

During the study, only the student investigator, the CI and immediate supervisory team will have access to the final study dataset at the UOS. On completion of the study, the Yorkshire Ambulance Service will be responsible for formalising archive arrangements.

11 Dissemination Policy

11.1 Research outputs

11.1.1 Anticipated research products

- A toolkit that can support clinicians in deciding the likelihood of an avoidable attendance before transporting the patient. This can be built into an electronic patient care record and used to support clinical assessment and decision making. Outputs will be disseminated trough peer review publications, presentation at relevant conferences.
- A Systematic review of risk prediction modelling in the urgent and emergency care system.
- New knowledge on what clinical assessment variables contribute most to risk stratifying patient's on scene.
- New knowledge on using risk prediction modelling in the prehospital setting and the potential feasibility of more complicated models such as machine learning in assisting clinical decision making.
- New knowledge on how to manage patient expectations on-scene.
- New knowledge on how to influence/change a clinician's decision on-scene by presenting them with patient risk.
- Two anonymous linked datasets of clinical care prehospital and in hospital data that other projects can use for further analysis.
- Further research plans for refining the tool, implementation and testing under real world conditions such as a randomised controlled trial.

11.2 Dissemination strategy

11.2.1 Strategy for disseminating to public and patients

Public members will be invited to co-produce a video that will be created for the public. This will introduce the concept of not transporting all patients and using a tool to help clinicians with making the decision to transport or not. They will also be invited to co-author conference abstracts and present findings with the researcher. As the research progresses, a PI group formed out of the Sheffield Emergency Care Forum (SECF) and interested members of the RDS PI event will steer and develop further dissemination strategies.

11.2.2 Strategy for disseminating to NHS

The research will be presented to the Association of Ambulance Chief Executives. This will be to highlight the findings and scope feasibility to implement nationwide. Contacts made at this level will be followed up at regional and local level. Contact with the National Leads for Urgent and Emergency Care will be maintained throughout the project and feedback will be invited from them at various stages. In addition, workshops will be put on for NHS staff and the work will be

presented to the UECRT and the National Ambulance Commissioners Network (NEWS). A Lay executive summary will be produced in digital format to disseminate widely on stakeholders web pages.

11.2.3 Strategy for disseminating to wider population

The linked data used in the research uses the Systemized Nomenclature of Medical Clinical Terms (SNOMED-CT), which is an international recognised clinical terminology. This would allow the possibility of international integration. Contacts will be made at the PAIC conference for future studies. Reproducibility of the tool outside of the NHS would require a validation step including calibration and discrimination.

11.2.4 Strategy for implementation

Areas for further research will be identified and grant applications will be made to secure funding. Once the model has been developed, there will be collaboration with NHSD. They will be the main route to implementation. There will also be an opportunity to work with industry partners who own patient care software.

The clinical terminology that the variables are captured (SNOMED CT) is advantageous in implementing the tool as the National Information Board (NIB) has mandated all NHS organisations capture clinical information in this language.

11.2 Authorship eligibility guidelines and any intended use of professional writers

Eligibility for authorship of proposed publications will be determined in accordance with International Committee of Medical Journal Editors (ICMJE) authorship criteria. Professional writers will not be used.

11.3 Intellectual Property

All Intellectual Property (IP) rights remain with the sponsor unless the NIHR (funder) deems the sponsor to be inefficient on acting upon the utilisation of such IP. In such a case, the funder has the right to take ownership of the IP.

12 References

1. National Audit Office. NHS Ambulance Services. Audit Report. , NHS England; 2017.

2. NHS England. Daily hospital situation report - Number of patients arriving by ambulance and ambulance handover delays. Statistical bulletin. NHS Improvement data collection; 2018.

3. Cooney D, Millin M, Carter A, Lawner B, Nable J, Wallus H. Ambulance diversion and emergency department offload delay: resource document for the national association of EMS physicians position statement. Prehospital Emergency Care. 2011; 15(4): p. 555-60.

4. Cooney D, Wojicik S, Seth N, Vasisko C, Stimson K. Evaluation of ambulance offload delay at a university hospital emergency department. International Journal of Emergency Medicine. 2013; 6(1): p. 15.

5. Eckstein M, Chan L. The effect of emergency department crowding on paramedic ambulance availability. Annals of emergency medicine. 2004; 43(1): p. 100-5.

6. Eckstein M, Isaacs M, Slovis C, Kaufman B, Loflin J, O'Connor R. Facilitating EMS turnaround intervals at hospitals in the face of receiving facility overcrowding. Prehospital Emergency Care. 2005; 9(3): p. 267-78.

7. Segal E, Verter V, Colacone A, Afilalo M. The in-hospital interval: a description of EMT time spent in the emergency department. Prehospital emergency care. 2006; 10(3): p. 378-82.

8. House J. What's causing England's A&E crisis? Lancet. 2013; 382(9888): p. 195-6.

9. Kingswell C, Shaban R, Crilly J. The lived experiences of patients and ambulance ramping in a regional Australian emergency department: an interpretive phenomenology study. Australasian Emergency Nursing Journal. 2015; 18(4): p. 182-9.

10. Hitchcock M, Crilly J, Gillespie B, Chaboyer W, Tippett V, Lind J. The effects of ambulance ramping on emergency department legnth of stay and in-patient mortality. Australasian Emergency Nursing Journal. 2010; 13(1-2): p. 17-24.

11. Hammond E, Shaban R, Holzhauser K, Crilly J, Melton N, Tippett V. An exploratory study to examine the phenomenon and practice of ambulance ramping at hospitals within the Queensland Health Southern Districts and the Queensland ambulance service Brisbane: Queensland Health & Griffith University ; 2012.

12. Kingswell C, Shaban R, Crilly J. Concepts, antecedents and consequences of ambulance ramping in the emergency department: A scoping review. Australasian Emergency Nursing Journal. 2017; 20: p. 153-160.

13. Miles J, O'Keeffe C, Jacques R, Stone T, Mason S. 17 Exploring ambulance conveyances to the emergency department: a descriptive analysis of non-urgent transports. Emerg Med J. 2017; 34: p. A872-A873.

14. Lowy A, Kohler B, Nicholl J. Attendance at accident and emergency departments: Unnecessary or inapropriate? Journal of public health medicine. 1994; 16(2): p. 134-140.

15. O'Keeffe C, Mason S, Jacques R, Nicholl J. Characterising non-urgent users of the emergency department (ED): A retrospective analysis of routine ED data. PLOSone. 2018; 13(2).

16. Department of Health. Taking healthcare to the patient: Transforming NHS Ambulance Services. London:; 2005.

17. Urgent and Emergency Care Review Team. High quality care for all, now and for future generations: Transforming urgent and emergency care services in England - Urgent and Emergency Care Review. End of Phase 1 report. Leeds: NHS England, Medical; 2013. Report No.: (PGR) 00691-00694.

18. NHS England. Five year forward view. ; 2014.

19. UEC Review Team and ECIST. Safer, Faster, Better: good practice in delivering urgent and emergency care. A guide for local health and social care communities. ; 2015.

20. Service NWA. Paramedic pathfinder and community care pathways V3. Bolton:; 2014.

21. Newton M, Tunn E, Moses I, Ratcliffe D, Mackway-Jones K. Clinical navigation for beginners: the clinical utility and safety of the Paramedic Pathfinder. Emergency Medicine Journal. 2014; 31: p. e29-e34.

22. O'Hara R, Johnson M, Hirst E, Weyman A, Shaw D, Mortimer P, et al. A qualitative study of decision-making and safety in ambulance service transitions. Health Services Delivery Research. 2014; 2(56).

23. Halter M, Vernon S, Snooks H, Porter A, Close J, Moore F. Complexity of the decision-making process of ambulance staff for assessment and referral of older people who have fallen: a qualitative study. Emergency Medicine Journal. 2011; 28: p. 44-50.

24. Burrell L, Noble A, Ridsdale L. Decision-making by ambulance clinicians in London when managing patients with epilepsy: a qualitative study. Emergency Medicine Journal. 2013; 30: p. 236-240.

25. Al-Sulaiti M, Snooks H, Porter A. Non-conveyance of 999 callers: Early findings related subsequent health services callers. Emergency Medicine Journal. 2009; 26: p. 3-4.

26. Simpson P, Thomas R, Bendall J, Lord B, Lord S, Close J. 'Popping nanna back into bed' – a qualitative exploration of paramedic decision making when caring for older people who have fallen. BMC Health Services Research. 2017; 17: p. 299.

27. Hoikka M, Silfvast T, Ala-Kokko T. A high proportion of prehospital emergency patients are not transported by ambulance: a retrospective cohort study in Northern Finland. Acta Anaesthesiologica Scandinavica. 2017; 61: p. 549-556.

28. Brydges M, Spearen C, Birze A, Tavares W. A Culture in Transition: Paramedic Experiences with Community Referral Programs. Canadian Journal of Emergency Medicine. 2015; 17(6): p. 631- 638.

29. Ambulances-in-zicht 2015. Ambulances-in-zicht 2015. [Online].; 2015 [cited 2017. Available from: http://www.vrgz.nl/nieuws/ambulances-in-zicht-2015/.

30. Ebben R, Vloet L, Speijers R, Tonjes N, Loef J, Pelgrim T, et al. A patient safety and professional perspective on non-conveyance in ambulance care: A systematic review. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2017; 25(71).

31. Fraess-Philips A. Can paramedics safely refuse transport of non-urgent patients? Prehospital and Disaster Medicine. 2016; 31(6): p. 667-674.

32. Tohira H, Williams T, Jacobs I, Bremner A, Finn J. The impact of new prehospital practitioners on ambulance transport to the emergency department: a systematic review and meta-analysis. Emergency Medicine Journal. 2014; 31: p. 88-94.

33. Mikolaizak A, Simpson P, Tiedemann A, Lord S, Close J. Systematic review of non-transportation rates and outcomes for older people who have fallen after ambulance service call-out. Australasian Journal of Ageing. 2013; 32(3): p. 147-157.

34. Turner J, Bjarkoy M, Coleman P, Goodacre S, Knowles E, Mason S. Building the evidence base in prehospital urgent and emergency care: Review 2. Evidence review. , Department of Health; 2010.

35. NHS England. Ambulance Quality Indicators. [Online].; 2017 [cited 2018. Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/ambulance-quality-indicators/.

36. Thompson M, Kaplan C, Cao Y, Bazzoli G, Waters T. Reliability of 30 day readmission measures used in the hospital readmission reduction program. Health Services Research. 2016; 51(6): p. 2095-2114.

37. Patton G, Thakore S. Reducing inappropriate emergency department attendances: a review of ambulance service attendances at a regional teaching hospital in Scotland. Emergency Medicine Journal. 2012; 30: p. 459-461.

38. Six A, Blackus B, Kelder J. Chest pain in the emergency room: value of the HEART score. Neth Heart J. 2008; 16: p. 191-6.

39. Aujesky D, Roy P, Verschuren F. Outpatient versus inpatient treatment for patients with acure pulmonary embolism: an international, open label, randomised, non-inferiority trial. Lancet. 2011; 378: p. 41-48.

40. Stanley A, Ashley D, Dalton H. Outpatient management of patients with low-risk uppergastrointestinal haemorrhage: multicentre validation and prospective evaluation. Lancet. 2009; 373: p. 42-7.

41. Ala L, Mack J, Shaw R. Selecting ambulatory emergency care (AEC) patients from the medical emergency in-take: the derivation and validation of the Amb score. Clinical Medicine. 2012; 12: p. 420-6.

42. Cameron A, Jones D, Logan E, O'Keeffe C, Mason S, Lowe D. Comparison of Glasgow Admission Prediction Score and Amb Score in predicting need for inpatient care. Emergency Medicine Journal. 2018; 35: p. 247-251.

43. Aronsky D, Jones I, Raines B. An integrated computerized triage system in the emergency department. AMIA Annu Symp Proc. 2008;: p. 16-20.

44. Levin S, France D, Mayberry S. The effects of computerised triage on nurse work behaviour. AMIA Annu Symp Proc. 2006;: p. 1005.

45. Dong S, Bullard M, Meurer D. Reliability of computerized emergency triage. Academic Emergency Medicine. 2006; 13: p. 269-275.

46. Levin S, Toerper M, Hamrock E, Hinson J, Barnes S, Gardner H, et al. Machine-learning based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. Annals of emergency medicine. 2017.

47. NHS Digital. Delivering linked datasets to support health and care delivery and research. 2018;(April):1–14.

48. Royston P, Moons K, Altman D, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ. 2009; 338.

49. Altman D, Vergouwe Y, Royston P, Moons K. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009; 338.

50. Moons K, Altman D, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ. 2009; 338.

51. Harrell F, Lee K, Mark D. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistical Medicine. 1996; 15: p. 361-87.

52. Laupacis A, Sekar N, Stiell I. Clinical prediction rules. A review and suggested modifications of methodological standards. JAMA. 1997; 277: p. 488-94.

53. Peduzzi P, Concato J, Feinstein A, Holford T. The importance of events per independant variable in proportional hazards regression analysis II: Accuracy and precision of regression estimates. Journal of Clinical Epidemiology. 1995; 48: p. 1503-10.

54. Concato J, Peduzzi P, Holford T, Feinstein A. Importance of events per independant variable in proportional hazards regression analysis I: Background, goals, and general strategy. Journal of Clinical Epidemiology. 1995; 48: p. 1495-501.

55. Wynants L, Bouwmeester W, Moons K, Moerbeek M, Timmerman D, Van Huffel S, et al. A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. Journal of Clinical Epidemiology. 2015; 68: p. 1406-1414.

13. Appendix

13.1 - Amendment History

Amendment No.	Protocol version no.	Date issued	Author(s) of changes	Details of changes made