BMJ Open

Data quality and 30-day survival for out-of-hospital cardiac arrest in the UK out-of-hospital cardiac arrest registry: a data linkage study

Sangeerthana Rajagopal,[1,2] Scott J Booth,[1] Terry P Brown,[1] Chen Ji,[1] Claire Hawkes,[1] A Niroshan Siriwardena,[3] Kim Kirby,[4] Sarah Black,[4] Robert Spaight,[5] Imogen Gunson,[6] Samantha J Brace-McDonnell,[1,2] Gavin D Perkins,[1,2] on behalf of OHCAO collaborators

[1]Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry, UK
[2]Heart of England NHS Foundation Trust, Birmingham, UK
[3]School of Health and Social Care, University of Lincoln, Lincoln, UK
[4]South Western Ambulance Service NHS Foundation Trust, Exeter, UK
[5]East Midlands Ambulance Service NHS Trust, Nottingham, UK
[6]West Midlands Ambulance Service NHS Foundation Trust, Brierley Hill, UK

Correspondence to
Professor Gavin D Perkins;
g.d.perkins@warwick.ac.uk

ABSTRACT

Objectives The Out-of-Hospital Cardiac Arrest Outcomes (OHCAO) project aims to understand the epidemiology and outcomes of out-of-hospital cardiac arrest (OHCA) across the UK. This data linkage study is a subproject of OHCAO. The aim was to establish the feasibility of linking OHCAO data to National Health Service (NHS) patient demographic data and Office for National Statistics (ONS) date of death data held on the NHS Personal Demographics Service (PDS) database to improve OHCAO demographic data quality and enable analysis of 30-day survival from OHCA.

Design and setting Data were collected from 1 January 2014 to 31 December 2014 as part of a prospective, observational study of OHCA attended by 10 English NHS Ambulance Services. 28 729 OHCA cases had resuscitation attempted by Emergency Medical Services and were included in the study. Data linkage was carried out using a data linkage service provided by NHS Digital, a national provider of health-related data. To assess data linkage feasibility a random sample of 3120 cases was selected. The sample was securely transferred to NHS Digital to be matched using OHCAO patient demographic data to return previously missing demographic data and provide ONS date of death data.

Results A total of 2513 (80.5%) OHCAO cases were matched to patients in the NHS PDS database. Using the linkage process, missing demographic data were retrieved for 1636 (72.7%) out of 2249 OHCAO cases that had previously incomplete demographic data. Returned ONS date of death data allowed analysis of 30-day survival status. The results showed a 30-day survival rate of 9.3%, reducing unknown survival status from 46.1% to 8.5%.

Conclusions In this sample, data linkage between the OHCAO registry and NHS PDS database was shown to be feasible, improving demographic data quality and allowing analysis of 30-day survival status.

Strengths and limitations of this study

► Data points collected as part of the Out-of-Hospital Cardiac Arrest Outcomes (OHCAO) project were based on established Utstein guidelines.
► The quality of demographic data collected by the OHCAO project was first improved through a list cleaning and patient status service provided by NHS Digital.
► Following list cleaning, exact data matches with Office for National Statistics date of death data allowed calculation of 30-day survival status.
► Provision of National Health Service numbers from OHCAO and NHS Digital provides potential for following long-term survival outcomes in OHCA patients through data linkage.
► Improved data linkage is reliant on improved data capture of patient demographic data by ambulance services.

resuscitation attempted.[1 2] This group suffers significant mortality and morbidity,[3 4] and improving outcomes from OHCA remains a worldwide research priority.[5]

Collecting high-quality data is essential as this forms the basis of decisions that ultimately impact on changes in care and healthcare resource allocation. Since 2011, survival to hospital discharge rates for OHCA has been reported as part of the National Health Service (NHS) England Ambulance Quality Indicators (AQIs), with significant variation reported ranging from 2.2% to 12.0%.[6] Regional variation in survival rates has also been observed worldwide.[7–9] Lilford et al highlighted that an important source of variation in reporting outcomes can be traced to the quality of data that results are based on.[10] Collecting survival to discharge data in England is a challenging process for

INTRODUCTION

Every year in the UK there are around 60 000 out-of-hospital cardiac arrests (OHCAs) attended by Emergency Medical Services (EMS) of which approximately 28 000 have

ambulance services as it involves tracking the patient's survival status directly with hospital emergency departments, which is time consuming and can be hindered by governance issues.[11][12]

Data collected in international OHCA registries enable comparisons of OHCA epidemiology and outcomes across different EMS systems.[13–15] The Utstein guidelines provide a structured template for collecting data on OHCA processes to support such comparisons.[16] To facilitate ease of reporting, the updated Utstein guidelines recommend collecting either 30-day survival or survival to hospital discharge as a core outcome.[17] The research literature suggests most international registries are able to report either of these OHCA outcome measures.[13–15] A recent example is the EuReCa ONE study which aimed to benchmark OHCA incidence, process and outcomes across 27 European countries and reported a combined survival to discharge or 30-day survival rate which ranged between 1.1% and 30.8%.[15]

Data linkage methodology has increasingly been used in medical research to establish outcomes. It involves linking information together from different sources that belong to the same individual.[18] Data linkage has been used by regional and national OHCA databases to confirm survival status through linkage with mortality databases.[4][19][20] Data linkage can address missing data issues, providing a centralised, high-quality database for research and service appraisal with the potential to allow longitudinal surveillance of OHCA patients.

The Out-of-Hospital Cardiac Arrest Outcomes (OHCAO) project is funded by the Resuscitation Council (UK), British Heart Foundation, and managed by the University of Warwick. It is a prospective observational study investigating the epidemiology and outcomes of OHCA patients across the UK.[21][22] This paper presents a subproject of the OHCAO project aiming to establish the feasibility of linking OHCAO registry data to NHS patient demographic data and Office for National Statistics (ONS) mortality data through the NHS Digital list cleaning and patient status service.

METHODS
Setting
The OHCAO project established a national UK OHCA registry to collect process and outcome data to facilitate OHCA research and quality improvement. Detailed information about the OHCAO project is available in the study protocol.[21] The 10 English NHS ambulance services collecting data for the OHCAO project cover approximately 54 million people, equating to 99.7% of the England population and 83.9% of the UK population.[23] Data were collected from 1 January 2014 to 31 December 2014 on 28 729 patients suffering OHCA in whom resuscitation was attempted by statutory EMS (an incidence rate of 53.2 per 100 000 of the English population).[22] This figure was reached after excluding individuals who achieved return of spontaneous circulation (ROSC)

before arrival of EMS (n=1711) and where resuscitation was not attempted as per national guidelines[24] due to the presence of a do not attempt resuscitation order (n=387) or signs incompatible with life or where resuscitation attempts would be futile (n=5403).

Aims and objectives
The overall aim of this project was to investigate the feasibility of linking a sample of OHCAO 2014 data to NHS patient demographic data and ONS date of death data held on the NHS Personal Demographics Service (PDS) database, using the NHS Digital list cleaning and patient status service, to improve OHCAO demographic data quality and allow calculation of 30-day survival from OHCA. The objectives were to (1) assess the match rate of combinations of OHCAO patient demographic variables in the sample (NHS number, surname, forename, date of birth (DOB) and home postcode) for linking to the NHS PDS database through NHS Digital list cleaning; (2) assess improvements in the completeness of OHCAO patient demographic variables through NHS Digital list cleaning and (3) create a linked OHCAO and NHS PDS database allowing analysis of 30-day survival from OHCA.

OHCAO project data collection
Core and supplemental Utstein variables were collected encompassing demographic, system, process and outcome data.[16] Each ambulance service has its own methods for OHCA case ascertainment, for example, electronic searches of patient report form databases for diagnostic codes indicating cardiac arrest. A trained member of the ambulance service clinical audit team entered eligible cases into a cardiac arrest database, followed by data cleaning and verification processes. Survival to hospital discharge data was collected directly from hospitals by the clinical audit team if data-sharing protocols were in place. Each ambulance service uploaded its data via a secure server to the OHCAO registry which is stored at the University of Warwick.

OHCAO data sample
To assess feasibility while minimising costs associated with data linkage, the analysis presented here represents a 10.9% sample of the 2014 OHCAO data, comprising 3120 OHCA patients. To avoid selection bias, the sample was selected using simple random sampling and stratified by ambulance service.

OHCAO data linkage to ONS mortality data
OHCAO to NHS PDS data linkage approval was received after submitting an application to the NHS Digital Data Access Request Service; additional approval was obtained from ONS for the release of mortality data. OHCAO submitted 3120 cases to NHS Digital, via the NHS Digital secure transfer system, detailing the following patient demographic variables of varying completeness: NHS number, surname, forename, DOB and home postcode.

NHS Digital is the national provider of data relating to health and social care in England. OHCAO used

the NHS Digital list cleaning and patient status service. The list cleaning service was used to validate submitted demographic data to ensure accuracy and improve data linkage outcomes. Validation was achieved by NHS Digital matching submitted demographic variables to NHS patient demographic data held on the PDS database. The PDS database is a national electronic database containing NHS patient demographic information, including NHS number, name and address. For each matched case, NHS Digital was asked to provide OHCAO with the following patient demographic information: NHS number, surname, forename and home postcode. These data were used to improve the percentage of missing data for these variables in the OHCAO sample.

NHS Digital used both automatic and manual matching techniques, using a combination of deterministic and probabilistic data linkage methods.[25 26] In deterministic data linkage, it is decided a priori what combination of patient identifiers to match on (eg, NHS number and DOB) and only complete agreement between records is considered a match. In probabilistic data linkage, weights are assigned to different patient identifiers (based on their discriminatory power) to assess the probability that two records are a match.[27] Cases were initially submitted for automatic matching which used a decision tree algorithm to provide matches. A subset of cases that failed automatic matching were resubmitted for manual matching.

As part of the patient status service, NHS Digital was also able to provide a date of death for deceased patients. The date of death data was held in the NHS PDS database and was sourced from ONS mortality data. OHCAO required information on deaths from 1 January 2014 until 31 January 2015. This was used to calculate 30-day survival. Where no date of death was provided, the patient was categorised as alive.

ONS date of death data

ONS mortality data contain all deaths registered in England and Wales. When a person dies, a formal medical certificate of death is produced, usually by a doctor, and which includes date of death. There is then a legal requirement for the death to be registered with the Registrar of Births, Deaths and Marriages through the local register office. The registration is typically performed by a close relative. The certification, and subsequent registration, of death may be delayed if the death is referred to a coroner for investigation (eg, if cause of death is unknown). However, the majority of deaths in England and Wales are registered within 5 days of the death date.[28 29] ONS receives death data in electronic form directly from register offices. All data received are subject to both initial and routine data quality and validation processes and are collected in line with the Statistics and Registration Service Act 2007.[28]

Analysis

An analysis was conducted to assess how particular demographic data points enabled linkage with NHS PDS data and whether data linkage improved the completeness of patient demographic data. This was done descriptively with breakdowns of data linkage match rates for all combinations of the OHCAO demographic variables sent to NHS Digital for data linkage.

The combined linked dataset was analysed to investigate 30-day survival rates calculated by evaluating if patients were alive ≥30 calendar days from the EMS OHCA incident date. The analysis was carried out pre-linkage and post-linkage, illustrating linkage effects. Thirty-day survival was calculated using OHCAO data where there was a date of death or date discharged >30 days after the OHCA incident date. Where there was an OHCAO date of death ≤30 days after the OHCA incident date or further ambulance service data indicating the patient was deceased on the day of the OHCA incident date (eg, hospital code indicating patient deceased and not conveyed to hospital), the patient was categorised as not surviving to 30 days. All other cases were categorised as unknown for patient's 30-day survival status. For the combined linked dataset, cases that were linked to ONS mortality data were categorised as 30-day survival where there was no date of death or where a date of death was provided that was >30 days after the OHCA incident date. Where there was an ONS date of death ≤30 days after the OHCA incident date, the patient was categorised as not surviving to 30 days. Where there was a contradiction in patient survival status between OHCAO data and ONS mortality data, then ONS mortality data superseded OHCAO data.

RESULTS

OHCAO data cleaning process

Of the 3120 cases transferred to NHS Digital, 2070 (66.3%) were automatically matched by the NHS Digital list cleaning algorithm while 1050 (33.7%) were not (figure 1). In total, 620 (19.9%) cases failing automatic matching were resubmitted for manual matching following which 437 (14.0%) were returned having been manually matched. Also, 430 cases (13.8%) were not resubmitted for manual matching as there was little chance of a match due to missing data points (252 cases only had 1 data point out of surname, forename, DOB and home postcode and 178 cases did not have any data points). Overall, 2513 (80.5%) cases were matched of which 7 (0.2%) cases could not be released due to the patient being lost to follow-up (one case, reason unknown) or the patient had registered a type 2 opt-out with NHS Digital, meaning that the patient's personal confidential data could not be released by NHS Digital for reasons other than their own direct care (six cases). In total, 607 (19.5%) cases could not be matched due to insufficient data for matching.

Data points required for matching through NHS Digital list cleaning

The percentage of each available demographic data point in the random sample of 3120 cases was similar to the
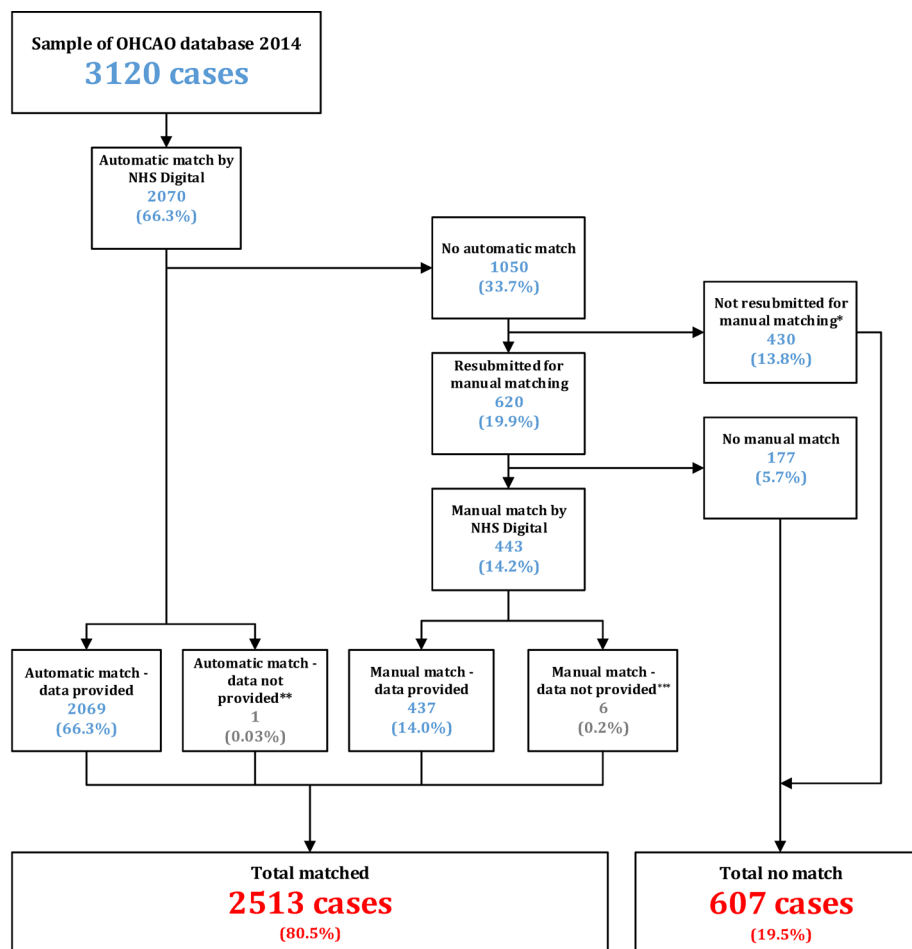
Figure 1 Data matching process linking Out-of-Hospital Cardiac Arrest Outcomes (OHCAO) data with National Health Service (NHS) Personal Demographics Service (PDS) data through the NHS Digital list cleaning and patient status service.

percentage of each available demographic data point in all 28 729 cases for 2014 (table 1). The data point determining the highest match rate was NHS number. One hundred per cent of cases with an NHS number were matched to the PDS database and therefore matched

to ONS mortality data with 99% of cases with an NHS number being automatically matched. However, only 31.7% of OHCAO cases had an NHS number.

Approximately a quarter (27.8%) of the sample had all 5 data points allowing a match to NHS PDS data (table 2).

Table 1 Total cases with each demographic data point collected by Out-of-Hospital Cardiac Arrest Outcomes (OHCAO) project

		National Health Service number	Surname	Forename	Date of birth	Postcode
Total OHCAO cases with data point (% of total 28 729 cases)		9510 (33.1%)	24 814 (86.4%)	24 686 (85.9%)	24 956 (86.9%)	15 017 (52.3%)
Total OHCAO sample cases with data point (% of total 3120 cases)		989 (31.7%)	2699 (86.5%)	2693 (86.3%)	2700 (86.5%)	1626 (52.1%)
Match status (% of sample cases with specified data point)	Total matched	989 (100%)	2506 (92.8%)	2505 (93.0%)	2408 (89.2%)	1566 (96.3%)
	Auto match	979 (99.0%)	2070 (76.7%)	2070 (76.9%)	2070 (76.6%)	1364 (83.9%)
	Manual match	10 (1.0%)	436 (16.2%)	435 (16.2%)	338 (12.5%)	202 (12.4%)
	No match	0 (0.0%)	193 (7.2%)	188 (7.0%)	292 (10.8%)	60 (3.7%)

Table 2 Combinations of data points required for linkage to the National Health Service (NHS) Personal Demographics Service database

Data points, n (n cases, % of total)	Combinations	n (total 3120 cases)	
		Matched, 2513 (n, % of total in data point category)	Unmatched, 607 (n, % of total in data point category)
5 (868, 27.8%)	NHS+surname+forename+DOB+postcode	868 (100%)	0
4 (815, 26.1%)	NHS+surname+DOB+postcode	0	0
	NHS+forename+DOB+postcode	0	0
	NHS+surname+forename+postcode	3 (0.4%)	0
	NHS+surname+forename+DOB	112 (13.7%)	0
	Surname+forename+DOB+postcode	665 (81.6%)	35 (4.3%)
	Total	780 (95.7%)	35 (4.3%)
3 (846, 27.1%)	NHS+surname+forename	1 (0.1%)	0
	NHS+surname+DOB	0	0
	NHS+surname+postcode	0	0
	NHS+forename+DOB	0	0
	NHS+forename+postcode	0	0
	NHS+DOB+postcode	0	0
	Surname+forename+DOB	760 (89.8%)	44 (5.2%)
	Surname+forename+postcode	27 (3.2%)	11 (1.3%)
	Surname+DOB+postcode	1 (0.1%)	1 (0.1%)
	Forename+DOB+postcode	1 (0.1%)	0
	Total	790 (93.4%)	56 (6.6%)
2 (156, 5.0%)	NHS+surname	0	0
	NHS+forename	1 (0.6%)	0
	NHS+DOB	0	0
	NHS+postcode	0	0
	Surname+forename	67 (42.9%)	82 (52.6%)
	Surname+DOB	0	0
	Surname+postcode	0	2 (1.3%)
	Forename+DOB	0	1 (0.6%)
	Forename+postcode	0	1 (0.6%)
	DOB+postcode	1 (0.6%)	1 (0.6%)
	Total	69 (44.2%)	87 (55.8%)
1 (257, 8.2%)	NHS	4 (1.6%)	0
	Surname	2 (0.8%)	18 (7.0%)
	Forename	0	14 (5.4%)
	DOB	0	210 (81.7%)
	Postcode	0	9 (3.5%)
	Total	6 (2.3%)	251 (97.7%)
0 (178, 5.7%)	Nil	0	178 (100%)

DOB, date of birth.

53.2% had 3–4 data points of which 93.4% and 95.7% were matched, respectively. Of these, all with NHS numbers were matched, while a combination of data points surname+forename+DOB+postcode and surname+forename+DOB resulted in match rates of 81.6% and 89.8%. However, cases where only 1 or 2 data points were

Table 3 Number of data points added to Out-of-Hospital Cardiac Arrest Outcomes data after NHS Digital list cleaning

Number of demographic data points increased by list cleaning	Cases (n)
0	1484 (47.6%)
1	804 (25.8%)
2	825 (26.4%)
3	7 (0.2%)

Table 4 Comparison of 30-day survival calculation pre-data and post-data linkage

30-day survival

Dataset 1: OHCAO data		Dataset 2: linked OHCAO and ONS data	
Yes	12 (0.4%)	Yes	290 (9.3%)
No	1670 (53.5%)	No	2566 (82.2%)
Unknown	1438 (46.1%)	Unknown	264 (8.5%)
Total	3120 (100%)	Total	3120 (100%)

OHCAO, Out-of-Hospital Cardiac Arrest Outcomes; ONS, Office for National Statistics.

provided were less likely to be matched (2.3% and 44.2%, respectively). In total, 178 (5.7%) cases had no OHCAO demographic data and could not be matched.

Data improvements after NHS Digital list cleaning and provision of ONS date of death data
Demographic improvements
After case matching, NHS Digital returned demographic data (forename, surname, NHS number, home postcode) and ONS date of death if applicable. In total, 1484 (47.6%) cases were not improved for any demographic data points (table 3). These cases were those where complete demographic data were already collected by OHCAO (868 cases), matching failed (607 cases) or data could not be released by NHS Digital due to the patient either being lost to follow-up or registering a type 2 opt-out with NHS Digital (six of the seven cases). All demographic data were already collected by OHCAO for 1 case out of these 7 cases and therefore were included in the aforementioned 868 cases. Lastly, for three cases OHCAO collected NHS+surname+forename+postcode and therefore these effectively could not be improved by matching as NHS Digital was not asked to provide DOB.

Of the 2249 cases with missing data, 1636 (72.7%) cases had demographic improvements following linkage. A quarter (25.8%) were improved by 1 demographic data point and a further quarter (26.4%) by 2 data points. Of the seven that were improved by 3 data points (table 3), OHCAO provided NHS number for four cases, surname for two and DOB and postcode for one.

NHS Digital returned NHS numbers for 1518 (48.7%) cases in which it was not already collected by OHCAO (see online supplementary table 1). OHCAO had already collected forename and surname in most cases (86.5% and 96.3%, respectively) which were least improved following matching.

Survival data improvements following provision of ONS date of death data
Thirty-day survival status (yes or no) using OHCAO data was confirmed for 1682 (53.9%) cases (table 4). Thirty-day survival was confirmed using OHCAO data if ambulance services provided a date of death or discharge date over 30 days after the OHCA incident date. Linking to ONS mortality data resulted in calculation of 30-day survival status (yes or no) for 2856 (91.5%) cases, a 37.6% improvement in 30-day survival status confirmation. The

pre-linkage 30-day survival rate was calculated as 0.4% and post-linkage as 9.3%.

Accuracy of OHCAO date of death data
In this sample, OHCAO reported a date of death for 1178 (37.8%) cases from ambulance services. In seven (0.6%) cases, death was not recorded with the ONS at the time of linkage. Of the 1942 (62.2%) cases where OHCAO could not confirm a date of death, 248 (12.8%) were recorded as alive at the time of linkage and 1137 (58.5%) had died according to ONS mortality data (table 5).

DISCUSSION
This study demonstrates the feasibility of linking OHCAO data to NHS patient demographic data and ONS date of death data through NHS Digital. In this sample of 3120 OHCAO cases, an 80.5% match rate was achieved and this enabled provision of registered death dates to calculate 30-day survival status. The results showed a 30-day survival rate of 9.3%, reducing unknown survival status from 46.1% to 8.5% (table 4). Additionally, demographic data quality improved for 1636 (52.4%) cases, with NHS numbers being provided for 1518 (48.7%) cases and postcodes for 942 (30.2%) cases where these data were missing in the OHCAO database.

Table 5 Comparison of date of death confirmed by Out-of-Hospital Cardiac Arrest Outcomes (OHCAO) and Office for National Statistics (ONS) data

		ONS confirmed survival status			Total (n, % total 3120 cases)
		Dead	Alive	No	
OHCAO project date of death provided	Yes	1114	7	57	1178 (37.8%)
	No	1137	248	557	1942 (62.2%)
Total (n, % total 3120 cases)		2251 (72.1%)	255 (8.2%)	614* (19.7%)	3120 (100%)

*Includes seven matched cases where data not provided by NHS Digital (one patient lost to follow-up (reason unknown), six patient registration of type 2 opt-out with NHS Digital).

The variability of cardiac arrest survival across ambulance services in England has been previously highlighted.[6] Where data from ambulance services do not follow a standard procedure, data collection variability may have significant effects on data quality and comparability between services. Increasingly, core outcome sets for specific research areas are developed outlining minimum datasets for routine collection and create a level of standardisation to compare studies and allow formation of meta-analyses.[30] In the field of OHCA, the Utstein guidelines have been developed.[16 17] However, a study investigating the level of missing data within primary outcomes in 283 Cochrane Reviews of all areas of clinical practice found that >50% of patient data were missing in 18% of reviews.[31] Furthermore, an analysis of 12 international OHCA registries collection of data using Utstein templates found that although all registries collected core variables, there were differences in interpretation of the template and recorded 'unknown' for a mean of 4.8 variables and 'missing' for 1.9 variables.[32] Therefore, minimum datasets are not sufficiently effective in reducing missing data.

The best data point provided by ambulance services to identify cases in the UK is the NHS number. It provides a unique identifier to resolve missing demographic data issues if no other demographic data are provided. One hundred per cent of OHCAO cases with an NHS number were matched to NHS PDS data; however, it was only available in a third (31.7%) of cases (table 1). Logistical difficulties exist in ascertaining an NHS number as it may not be available in the out-of-hospital setting. However, this study found that providing at least 3–4 demographic variables other than an NHS number resulted in a match rate of up to 89.8%, depending on the combination and especially if forename and surname were provided. This also allowed provision of an NHS number in 48.7% of cases where it was not collected by the OHCAO project (see online supplementary table 1). If less than a threshold of 3 data points were provided, this study found a lower potential for matching (0%–44.2%, table 2). Our findings support previous research showing that the ability to successfully link international OHCA databases to outcome data is dependent on the provision and completeness of patient identifiers. For example, the Danish Cardiac Arrest Registry was able to link to the Danish Civil Registration System to confirm 30-day survival for 100% of OHCA patients due to 100% provision of a unique Civil Registration Number.[20] Conversely, a study from the USA showed limited feasibility for linking OHCA patients to longitudinal outcomes when there was no unique patient identifiers available and there was variability in completeness of patient demographic data, resulting in a linkage rate of only 34.2%.[33]

NHS Digital list cleaning increased the number of OHCAO cases with a validated NHS number by 1518 (48.7%) to 2507 (80.4%) cases, suggesting that data linkage is a feasible method for linking an OHCA dataset to the national mortality dataset. The current process for ambulance services in England to confirm survival to discharge from OHCA is challenging,[11 12] and using the NHS Digital list cleaning and patient status service to calculate 30-day survival from OHCA may be a viable alternative. The results of this study also suggest the potential to use data linkage for further avenues of research relating to OHCA in the UK. Data linkage can be used to follow OHCA patients longitudinally, for example, to investigate predictors of survival at 1 year, 5 years and beyond.[4 34] Furthermore, data linkage can be used to evaluate the complete patient care pathway by linking to existing routinely collected hospital data sources. For example, hospital interventions and hospital length of stay via Hospital Episode Statistics and intensive care interventions via the Intensive Care National Audit and Research Centre. NHS Digital also provided postcodes for a further 942 (30.2%) cases, which increases the potential to examine the influence of neighbourhood characteristics, such as population density and social deprivation, on OHCA incidence, whether an event is witnessed, and if they receive bystander cardiopulmonary resuscitation (CPR).[35 36]

The OHCAO project was able to collect a date of death for 1178 (37.8%) cases. Interestingly, a date of death was not recorded with the ONS for seven of these cases, indicating that the patients were still alive at the time of linkage. Such errors may lead to incorrect reporting of cardiac arrest survival as part of the NHS England AQIs. This is an important finding as this shows the importance of data linkage to correct database errors.

This study's strengths lie in its standardised procedures for OHCA case definition and data collection, with the data points collected based on established Utstein guidelines.[16] A further strength is that NHS Digital used both deterministic and probabilistic data linkage methods; they have different strengths and using both methods may enhance linkage performance.[37] Deterministic linkage methods have greater specificity but require exact matches between records, while probabilistic data linkage has greater sensitivity, working better with poorer quality data as it allows imperfect matches between records.[27] For example, the returned demographic data for the linked cases showed that 14 OHCAO cases with between 4 and 5 data points were linked despite having an erroneous NHS number. This allowed correction of the inaccurate NHS number in the OHCAO sample. Finally, successful data linkage enabled access to high-quality national date of death data from ONS that is subject to rigorous data quality and validation processes.[28]

Limitations

This study had several limitations. First, only 868 (27.8%) cases had all 5 OHCAO data points, while 178 (5.7%) cases had missing data for all OHCAO data points. Missing data is an issue in OHCA registries,[32] and improved data linkage in the OHCAO project is reliant on improved data capture of patient demographic data by ambulance services. While NHS numbers were provided for only 989 (31.7%) OHCAO cases, one ambulance service provided

NHS numbers for 100% of their cases. This suggests potential for the OHCAO project to work with ambulance services to increase provision of patient demographic data to improve data linkage. Second, following linkage 30-day survival status remained unknown for 264 (8.5%) cases. Data not missing completely at random can bias results.[38] For example, if those 264 patients survived to 30 days the overall 30-day survival rate would be 17.8% (584 cases) instead of 9.3% (290 cases). Third, where no date of death was provided, cases were categorised as alive. However, absence of recorded death may mean registration of death has been delayed, for example, due to a coroner's inquest. Although it should also be noted that NHS Digital did not commence data linkage until >12 months (March 2016) after the date (31 January 2015) where 30-day survival could be calculated for patients in the sample suffering an OHCA on 31 December 2014. ONS data for 2014–2015 show that only 6.1% of deaths in England and Wales required a coroner's inquest[28] and the average time of an inquest was 24 weeks.[39 40] Furthermore, ONS data from 2011 report that overall 94% of deaths were registered within 1 month.[29] Finally, where the quality and completeness of data is variable data linkage errors can occur, which can bias reported outcomes.[41] Deterministic data linkage methods increase the likelihood of false negative matches (not matching to a correct match), while probabilistic data linkage increases the likelihood of false positive matches (matching to an incorrect match).[27] To quantify how data linkage errors may impact on study findings and outcomes, a formal data linkage validation evaluation is required.[18] This was beyond the scope of this study but should be conducted if OHCAO establishes a data linkage programme.

CONCLUSIONS

This study shows the feasibility of linking data from the UK OHCAO project to NHS patient demographic and ONS date of death data using the NHS Digital list cleaning and patient status service. This enabled analysis of 30-day survival status which may be of use to the NHS in terms of resource planning and directing service provision. Missing NHS numbers are a significant obstacle to successful data linkage, and this study found that if at least forename and surname are collected with one other demographic data point, there is a high chance of retrieving missing NHS numbers. Demographic data were improved for over half of cases and can be used as a means of creating a registry of OHCA patients to investigate postresuscitation care and longitudinal outcomes.

REFERENCES

1. *Resuscitation Council (UK), NHS England, British Heart Foundation*. Consensus paper on out-of-hospital cardiac arrest in England: Resuscitation Council, 2014.
2. Ambulance Service Association. *National cardiac arrest audit report*. London: Ambulance Service Association, 2006.
3. Sasson C, Rogers MA, Dahl J, et al. Predictors of survival from out-of-hospital cardiac arrest: a systematic review and meta-analysis. *Circ Cardiovasc Qual Outcomes* 2010;3:63–81.
4. Shuvy M, Morrison LJ, Koh M, et al. Long-term clinical outcomes and predictors for survivors of out-of-hospital cardiac arrest. *Resuscitation* 2017;112:59–64.
5. Institute of Medicine Committee on the Treatment of Cardiac Arrest. *Strategies to improve cardiac arrest survival. A time to act*: National Academies Press, 2015.
6. Perkins GD, Cooke MW. Variability in cardiac arrest survival: the NHS ambulance service quality indicators. *Emerg Med J* 2012;29:3–5.
7. Girotra S, van Diepen S, Nallamothu BK, et al. Regional variation in out-of-hospital cardiac arrest survival in the United States. *Circulation* 2016;133:2159–68.
8. Okubo M, Kiyohara K, Iwami T, et al. Nationwide and regional trends in survival from out-of-hospital cardiac arrest in Japan: A 10-year cohort study from 2005 to 2014. *Resuscitation* 2017;115:120–8.
9. Strömsöe A, Svensson L, Axelsson ÅB, et al. Improved outcome in Sweden after out-of-hospital cardiac arrest and possible association with improvements in every link in the chain of survival. *Eur Heart J* 2015;36:863–71.
10. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004;363:1147–54.
11. Fothergill R, Brace-McDonnell SJ, Perkins GD. Variation in epidemiology and outcomes from cardiac arrest. *Resuscitation* 2014;85:1610–1.
12. Perkins GD, Lall R, Quinn T, et al. Mechanical versus manual chest compression for out-of-hospital cardiac arrest (PARAMEDIC):

a pragmatic, cluster randomised controlled trial. *Lancet* 2015;385:947–55.

13. Berdowski J, Berg RA, Tijssen JG, *et al*. Global incidences of out-of-hospital cardiac arrest and survival rates: Systematic review of 67 prospective studies. *Resuscitation* 2010;81:1479–87.

14. Ong ME, Shin SD, De Souza NN, *et al*. Outcomes for out-of-hospital cardiac arrests across 7 countries in Asia: The Pan Asian Resuscitation Outcomes Study (PAROS). *Resuscitation* 2015;96:100–8.

15. Gräsner JT, Lefering R, Koster RW, *et al*. EuReCa ONE-27 Nations, ONE Europe, ONE Registry: a prospective one month analysis of out-of-hospital cardiac arrest outcomes in 27 countries in Europe. *Resuscitation* 2016;105:188–95.

16. Jacobs I, Nadkarni V, Bahr J, *et al*. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update and simplification of the Utstein templates for resuscitation registries. *Resuscitation* 2004;63:233–49.

17. Perkins GD, Jacobs IG, Nadkarni VM, *et al*. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update of the utstein resuscitation registry templates for out-of-hospital cardiac arrest. *Resuscitation* 2015;96:328–40.

18. Moore CL, Amin J, Gidding HF, *et al*. A new method for assessing how sensitivity and specificity of linkage studies affects estimation. *PLoS One* 2014;9:e103690–8.

19. Elmer J, Rittenberger JC, Coppler PJ, *et al*. Long-term survival benefit from treatment at a specialty center after cardiac arrest. *Resuscitation* 2016;108:48–53.

20. Hamilton A, Steinmetz J, Wissenberg M, *et al*. Association between prehospital physician involvement and survival after out-of-hospital cardiac arrest: A Danish nationwide observational study. *Resuscitation* 2016;108:95–101.

21. Perkins GD, Brace-McDonnell SJ. The UK Out of Hospital Cardiac Arrest Outcome (OHCAO) project. *BMJ Open* 2015;5:e008736.

22. Hawkes C, Booth S, Ji C, *et al*. Epidemiology and outcomes from out-of-hospital cardiac arrests in England. *Resuscitation* 2017;110:133–40.

23. Office for National Statistics Annual Mid-Year Population Estimates. 2014 http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/2015-06-25(Accessed 7 July 2017).

24. Joint Royal Colleges Ambulance Liaison Committee and Association of Ambulance Chief Executives. *UK Ambulance Services Clinical Practice Guidelines 2016*. Bridgwater: Class Professional Publishing, 2016.

25. Jaro MA. Probabilistic linkage of large public health data files. *Stat Med* 1995;14:491–8.

26. Méray N, Reitsma JB, Ravelli AC, *et al*. Probabilistic record linkage is a valid and transparent tool to combine databases without a patient identification number. *J Clin Epidemiol* 2007;60:883.e1–883.e11.

27. Zhu Y, Matsuyama Y, Ohashi Y, *et al*. When to conduct probabilistic linkage vs. deterministic linkage? A simulation study. *J Biomed Inform* 2015;56:80–6.

28. Office for National Statistics User guide to mortality statistics. 2017 https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsjuly2017

29. Office for National Statistics Impact of registration delays on mortality statistics. 2011 http://webarchive.nationalarchives.gov.uk/20160106020016/http://www.ons.gov.uk/ons/guide-method/user-guidance/health-and-life-events/impact-of-registration-delays-on-mortality-statistics/index.html

30. Kirkham JJ, Gorst S, Altman DG, *et al*. Core Outcome Set-STAndards for Reporting: The COS-STAR Statement. *PLoS Med* 2016;13:e1002148.

31. Kirkham JJ, Gargon E, Clarke M, *et al*. Can a core outcome set improve the quality of systematic reviews?--a survey of the Co-ordinating Editors of Cochrane Review Groups. *Trials* 2013;14:21.

32. Nishiyama C, Brown SP, May S, *et al*. Apples to apples or apples to oranges? International variation in reporting of process and outcome of care for out-of-hospital cardiac arrest. *Resuscitation* 2014;85:1599–609.

33. Mumma BE, Diercks DB, Danielsen B, *et al*. Probabilistic Linkage of Prehospital and Outcomes Data in Out-of-hospital Cardiac Arrest. *Prehosp Emerg Care* 2015;19:358–64.

34. Dumas F, Rea TD. Long-term prognosis following resuscitation from out-of-hospital cardiac arrest: role of aetiology and presenting arrest rhythm. *Resuscitation* 2012;83:1001–5.

35. Fosbøl EL, Dupre ME, Strauss B, *et al*. Association of neighborhood characteristics with incidence of out-of-hospital cardiac arrest and rates of bystander-initiated CPR: implications for community-based education intervention. *Resuscitation* 2014;85:1512–7.

36. Moon S, Bobrow BJ, Vadeboncoeur TF, *et al*. Disparities in bystander CPR provision and survival from out-of-hospital cardiac arrest according to neighborhood ethnicity. *Am J Emerg Med* 2014;32:1041–5.

37. Gomatam S, Carter R, Ariet M, *et al*. An empirical comparison of record linkage procedures. *Stat Med* 2002;21:1485–96.

38. He Y. Missing data analysis using multiple imputation. *Circ Cardiovasc Qual Outcomes* 2010;3:98–105.

39. Ministry of Justice Coroners Statistics 2014 England and Wales. 2015 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427720/coroners-statistics-2014.pdf

40. Ministry of Justice Coroners Statistics 2015 England and Wales. 2016 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/607728/coroners-statistics-2015.pdf

41. Hagger-Johnson G, Harron K, Fleming T, *et al*. Data linkage errors in hospital administrative data when applying a pseudonymisation algorithm to paediatric intensive care records. *BMJ Open* 2015;5:e008118.