Challenges of Using Probabilistic Linkage Methodology to Characterize Post-Cardiac Arrest Care in Michigan

Robert Swor, Lihua Qu, Kevin Putman, Kelly N. Sawyer, Robert Domeier, Jennifer Fowler & William Fales

To cite this article: Robert Swor, Lihua Qu, Kevin Putman, Kelly N. Sawyer, Robert Domeier, Jennifer Fowler & William Fales (2017): Challenges of Using Probabilistic Linkage Methodology to Characterize Post-Cardiac Arrest Care in Michigan, Prehospital Emergency Care, DOI: 10.1080/10903127.2017.1362086

To link to this article: https://doi.org/10.1080/10903127.2017.1362086

Published online: 14 Sep 2017.

Submit your article to this journal

Article views: 37

View related articles

View Crossmark data
CHALLENGES OF USING PROBABILISTIC LINKAGE METHODOLOGY TO CHARACTERIZE POST-CARDIAC ARREST CARE IN MICHIGAN

Robert Swor, DO, Lihua Qu, MD, Kevin Putman, MS, Kelly N. Sawyer, MD, Robert Domeier, MD, Jennifer Fowler, RN, William Fales, MD

ABSTRACT

**Background:** To improve survival of patients resuscitated from out of hospital cardiac arrest (OCHA), data is needed to assess and improve inpatient post-resuscitation care. Our objective was to apply probabilistic linkage methodology to link EMS and inpatient databases and evaluate whether it may be used to describe post-arrest care in Michigan. **Methods:** We performed a retrospective study to describe post-cardiac arrest care in adult OHCA patients who were transported to Michigan hospitals from July 1, 2010, to June 30, 2013. Using probabilistic linkage methodology we linked two databases, the Michigan EMS Information System (MI_EMSIS) and the Michigan Inpatient Database (MIDB), which describes inpatient care and outcome of all admissions. Rates of case incidence and survival were compared to published literature. We compared the linked dataset to existing cardiac arrest databases from three counties to evaluate the quality of this linkage. **Results:** Multiple iterations of match strategies were used to create a linked EMS-inpatient dataset. There were 12,838 MI_EMSIS cardiac arrest records of which 1,977 were matched with MIDB records, identifying them as surviving to hospital admission. Of these 590 (30.0%) survived to hospital discharge. The annual survival incidence/100,000 population to admission was 6.93/100,000 and survival incidence to discharge was 2.1/100,000. The matched dataset was compared to county databases identified a limited sensitivity [48.2%, 95% CI 42.1%–55.3%] and positive predictive value [64.4%, 95% CI 56.8%–71.3%]. **Conclusion:** Use of the MI_EMSISEMS database and the Michigan Inpatient database was feasible and produced rates of cardiac arrest admission and survival rates similar to published literature. This process yielded a limited match compared to existing county cardiac arrest databases. We conclude that such a linked dataset is useful for descriptive purposes but not as a population based dataset to evaluate statewide post-cardiac arrest care. **Key words:** cardiac arrest; outcome; linkage; emergency medical services

PREHOSPITAL EMERGENCY CARE 2017; Early Online:1–6

INTRODUCTION

Research efforts and published literature have focused on interventions that improve cardiac arrest survival in the community.1–3 Most of these interventions (e.g., community CPR provision, CPR quality metrics, early defibrillation) have addressed care rendered outside of the hospital. Attention to in-hospital interventions, such as post-arrest targeted temperature management (TTM), immediate percutaneous coronary intervention (PCI), and more aggressive post-arrest neuro-critical care, have occurred more recently.4–8 A number of authors and professional organizations have advocated for specialized “cardiac arrest centers” that have the commitment, capability and experience to provide optimal care for these patients.9–12 Data that links EMS field care and inpatient post-arrest care is needed.

Scant population-based data exists that describes the spectrum of post-resuscitation care. Two large, national databases have prospectively accumulated patient-level data (Resuscitations Outcome Consortium, ROC, and Cardiac Arrest to Enhance Outcomes, CARES),13,14 but they lack robust data on inpatient care and predominantly draw from EMS systems that have the capacity to participate in NIH sponsored clinical trials (ROC), or have leadership that is interested in the analysis and provision of OHCA care (CARES). Thus, large databases of cardiac arrest data are often biased by participation and may not reflect real-world implementation of post-cardiac arrest care guidelines.

In its recent report on strategies to improve cardiac arrest survival, the Institute of Medicine has identified
comprehensive surveillance and data reporting as the foundational element of improving outcome from cardiac arrest.\textsuperscript{15} The first recommendation of this report calls for “the development of a National Cardiac Arrest Registry to ... help increase federal and state accountability for current system performance and promote actions to improve cardiac arrest outcomes.”\textsuperscript{16} The development of such a registry will require the investment of a tremendous amount of resources and time. Until such a national registry is developed, strategies to utilize other types of “big data” are needed.

Probabilistic linkage of large administrative databases is an epidemiological tool that has been used to link EMS data with hospital processes and outcomes for trauma,\textsuperscript{16} burns,\textsuperscript{17} and stroke.\textsuperscript{18} This methodology allows matching of community level data with care provided in the health care system. The resultant dataset produces descriptive data of the care to these patients which is a crucial first step to quality improvement. These studies utilized unique identifiers to facilitate patient matching and yield high quality statewide databases. No literature using such identifiers exists for post-resuscitative care. One study has attempted to link EMS and inpatient databases for cardiac arrest care and identified fundamental problems with use of EMS data for this process.\textsuperscript{19}

For these reasons, our objective was to apply probabilistic linkage methodology to link EMS and inpatient databases and evaluate whether it may be used to describe post-arrest care in Michigan.

METHODS
We used two existing datasets—the Michigan Emergency Medical Services Information System (MI_EMSIS) and the Michigan Inpatient Database (MIDB)—to identify out of hospital cardiac arrest (OHCA) patients and their outcomes. The study period was from July 1, 2010, to June 30, 2013.

MI_EMSIS is based on the National EMS Information system (NEMSIS v 2.1.1), has been mandated in Michigan since 2009 and had almost 1.3 million out of hospital records submitted to it by 2012. During this study period, it received data from more than 80% of EMS agencies in the state of Michigan, covering more than 90% of the state’s population. EMS care is typical of large Midwestern states with a predominance of full time advanced life support municipal and private EMS services, with rural areas covered by volunteer basic life support services. Cardiac arrest treatment is provided using statewide EMS protocols, which have some local variation. Statewide protocols for resuscitated cardiac arrest patients call for transport to hospitals with interventional capability when practical. EMS data is submitted by EMS agencies across the state that use a number of different processes (single agency wide submission, submission by individual providers at the completion of each incident, manual data entry, computerized upload). The database is managed by a single data manager for the state, who is responsible for data quality management and recruitment. MI_EMSIS includes key demographic and process variables predictive of cardiac arrest survival, including Utstein style cardiac arrest data variables.\textsuperscript{20}

EMS cases were initially included if they had a primary or secondary impression of cardiac arrest, and were subsequently excluded if they had no cardiac resuscitation attempt, were traumatic in nature, were <19 years of age, had a do not resuscitate order, or had failed resuscitation on scene.

The MIDB is maintained by the Michigan Hospital Association and records diagnostic, procedural, and outcome data on all admissions to acute care hospitals in Michigan. Each hospital designates a registrar that reviews and assures data quality of submissions and addresses data queries. Data are subsequently reviewed by the Michigan Hospital Association data collection team. MIDB records patient outcome and disposition, as well as up to 20 diagnostic and procedural codes using ICD-9 CM codes during the study period.

Both databases are de-identified and collect patient demographics, temporal information regarding an event (incident or hospital admission date), and location information (incident city, county and receiving hospital). The study was approved by the institutional review boards of Beaumont Health System and the Michigan Department of Health and Human Services.

MATCHING PROCESS
We matched cases from MI_EMSIS, using incidents which had a provider primary or secondary impression of “cardiac arrest”. The MIDB was queried for all cases with a primary or admission diagnosis of cardiac arrest (ICD-9 = 427.5) or ventricular fibrillation (ICD-9 = 427.41). We used the PROC SQL function of SAS (v 9.4, Cary, NC.), using the method of Wright\textsuperscript{21} to match variables between two datasets. Match scores were calculated using a composite point score from points assigned to each variable match. Match variables included age, gender, date of event, receiving hospital, and county of original incident. Ethnicity was not used as a variable as it was missing in a large number of cases. Investigators assigned weighted scores based on perceived value of each variable to provide an accurate and unique match (gender (90 points), incident date +/- 1 year (85 points), age (80 points) receiving hospital (80 points), and incident county (75 points)). Distribution of total match scores were displayed graphically (Figure 1) and a cut point was determined visually to separate match and non-match pairs. Multiple iterations of match strategies were used to decide on an optimal cut point based on yield of match and rate of duplicate record generation. The final dataset was con-
FIGURE 1. Histogram of similarity scores for probabilistic linkage. Cut point for match. Histogram presents total score for matching of individual cases. Majority of cases have lower match scores and are less likely to represent single matched case in EMS and hospital database.

FIGURE 2. Patient flow diagram. *Other includes duplicate EMS responses removed manually, air medical cases, transport to out of state hospitals, DNR, and unknown destination.

RESULTS

During the study period there were 33,080 OHCA obtained from MI_EMSIS; 20,081 were excluded (Figure 2), leaving a total of 12,838 unique EMS transported cardiac arrest cases. There were 4,479 MIDB cases with admit diagnosis cardiac arrest and/or VF. Demographic data for these datasets as well as the gold standard dataset are presented in Table 1. The final matched dataset yielded 1,977 (15.4% of EMS incidents) cases. Although this would appear to represent a poor match, only EMS cases that survive to hospital admission would be matched, and this proportion of cases reflects the rate of hospital admission for EMS cases. For comparison in this state, the rate of hospital admission post-cardiac arrest has been reported as 23.4% (2014, CARES annual report. The survival incidence to admission was 6.93/100,000/population/year and survival incidence to discharge was 2.1/100,000/population/year.

The matched dataset was compared to three existing county databases and identified a limited sensitivity (48.2%, 95% CI 42.1%–55.3%) and positive predictive value (64.4%, 95% CI 56.8%–71.3%). There was substantial variability in sensitivity and positive predictive value rates by county (Table 2). There was no significant difference in percentage match by year (53.6%–64.4%, P = 0.35).
Table 1. Demographics and outcome characteristics of study datasets

<table>
<thead>
<tr>
<th></th>
<th>Initial EMS Dataset N = 12838</th>
<th>Hospital Dataset N = 4479</th>
<th>Final Matched Dataset N = 1977</th>
<th>Gold standard database* N = 277</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Male</td>
<td>7904 (61.6%)</td>
<td>2631 (58.6%)</td>
<td>1237 (62.7%)</td>
<td>279</td>
</tr>
<tr>
<td>Age (mean, range)</td>
<td>64.4 (18–113)</td>
<td>63.0 (18–104)</td>
<td>64.8 (+/- 15.8)</td>
<td>61.3 (20–97)</td>
</tr>
<tr>
<td>Ventricular Fibrillation (%)</td>
<td>2227 (17.3%)</td>
<td>1434 (32.0%)</td>
<td>23.5%</td>
<td></td>
</tr>
<tr>
<td>Bystander Witnessed (%)</td>
<td>4797 (37.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return of Spontaneous Circulation</td>
<td>3599 (28.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival to Hospital Admission (%)†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival to Hospital Discharge (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Datasets represent different points in time and characteristics and will have not have same variables in them, resulting in empty data fields.

*Gold standard databases include limited data from 3 counties with existing cardiac arrest database, and does not include clinical characteristics.

†Hospital Dataset, Matched dataset and Gold Standard database are 100% admissions.

Table 2. Comparison of linked vs. county post-arrest databases

<table>
<thead>
<tr>
<th>County</th>
<th>N</th>
<th>Sensitivity (%), 95% CI</th>
<th>Positive Predictive Value (%), 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>138</td>
<td>43.5% (35.0, 52.4)</td>
<td>89.1% (78.2, 95.1)</td>
</tr>
<tr>
<td>B</td>
<td>51</td>
<td>47.5% (26.4, 69.7)</td>
<td>25.0% (13.2, 41.5)</td>
</tr>
<tr>
<td>C</td>
<td>105</td>
<td>59.5% (47.8, 70.2)</td>
<td>64.4% (57.2, 75.0)</td>
</tr>
<tr>
<td>Total</td>
<td>294</td>
<td>48.9% (42.4, 55.5)</td>
<td>64.4% (56.8, 71.4)</td>
</tr>
</tbody>
</table>

DISCUSSION

We used a probabilistic linkage methodology to construct a database of patients that were resuscitated from OHCA. This dataset was comparable to, but lower than prior literature in the rate of survival to hospital admission (15.3% vs. 24.3%, CARES).13 Similarly, we identified a comparable but lower case incidence of admitted patients (6.9/100,000 population/year vs. 13.4/100,000/year, estimated from data presented by Nichol et al.).14 Finally, the sensitivity and positive predictive value of this matched database was disappointing when compared to existing regional datasets. This study identifies that probabilistic linkage is a feasible tool for linking EMS and inpatient databases for evaluating post-resuscitative care but also highlights the challenges in using a state EMS information system, which is similar to many that have been implemented nationally.22 Our first challenge was that our MI_EMSIS dataset did not have all cases that existed in our “gold standard communities,” despite both data sources receiving cases from the same EMS agencies. The MI_EMSIS dataset only had 202/234 (86.3%) of the cardiac arrest admissions found in the three county databases. This limited our ability to have a high quality match of MI_EMSIS-MIDB that would compare favorably to the gold standard. Part of this low rate of case ascertainment may be a result of our initial case definition from MI_EMSIS. Prior work by Mumma yielded a higher rate of cardiac arrest cases but utilized broader definitions of cardiac arrest, using cardiac arrest procedures or medications administered to identify cardiac arrest cases.19

More likely, we believe the lower incidence of cases reflects the quality of the datasets. Fundamentally, EMS data systems are not registries, similar to trauma registries or inpatient databases. EMS system databases are populated by frontline clinicians who have competing interests and whose primary focus is not provision of quality data. They are not uniformly educated on data collection and variation in basic items such as the definition of a cardiac arrest (e.g., resuscitation initiated and terminated, not initiated, CPR started by bystander but with a pulse on EMS arrival) complicates efforts to quantify care provided to post-arrest patients. The methodology for submission of data to a statewide EMS registry varies by local EMS system, with some using automated data upload, others providing direct data input. The sheer number of first responders, BLS providers, and paramedics assure substantial variation in data collection. Although the MI_EMSIS in Michigan has yielded high rates of data collection compared to other states (unpublished communication), we learned that not all large EMS systems in the state submitted data to MI_EMSIS. While data submission to the MI_EMSIS is mandated by the Michigan Department of Health and Human Services, assurance of compliance with data entry and problem solving of technical glitches require significant resources. Statewide submission of data to the MIDB was much better but also incomplete. During our data cleaning, we identified at least one small health system that did not identify any cardiac arrest patients.

The value of billing codes (ICD-9 CM during the study period) for identification of cardiac arrest cases has been questioned by Coppler et al.23 This paper evaluated whether billing codes could be used to identify cardiac arrest using emergency department discharge codes. They found a very low rate of match
to an externally managed cardiac arrest registry. Their billing (administrative) dataset, however, identified a population with a 90% mortality rate. This was substantially different from the survival rate of our inpatient dataset (38.3%). We also used MI EMSIS to identify cases and matched to hospital admission diagnosis, which is a different methodology. This work also used a broader range of ICD-9 codes to identify cardiac arrest, including diagnosis of ventricular tachycardia and respiratory arrest. An expanded definition of hospital codes may have increased our rate of match. Future investigation regarding the value of billing codes to match to EMS datasets would help advance this approach to analysis.

Our low rate of matching could also be due to statistical methodology. While this method has been previously used for other epidemiological purposes, using de-identified datasets whose data quality has not been validated decreases the likelihood of successful matching. States that have successfully used probabilistic linkage for emergency care purposes have used existing identifiers. We believe that an appropriate conclusion from our work is that EMS information systems should have unique identifiers that are recorded at hospitals to promote integration of data across the health care continuum.

LIMITATIONS

We believe this study highlights the potential utility and limitations of using a statewide NEMSIS EMS database to characterize inpatient care. Our low rates of case and survival incidence identify that MI EMSIS is not currently suitable as a population based dataset. First, we identified early in the project that some EMS systems, for a number of reasons, did not submit data to MI EMSIS. Our low rate of match compared to existing “gold standards” identify that, while MI EMSIS may be used to identify EMS patients beyond the field, matching with inpatient data did not yield a dataset comparable to existing local datasets. Our “gold standard” databases were variable in their matching characteristics compared to our final linked dataset, identifying them as an imperfect gold standard. No single community showed a highly sensitive rate of match. Overall, the low rate of match was, in part, because local cardiac arrest cases were not uniformly submitted to the MI EMSIS database. Future versions of the parent national NEMSIS dataset are being implemented, which will improve the granularity of EMS data; however, it will not, by itself, improve rates of initial data capture. Our study evaluates the use of a NEMSIS database in one state, which may not be reflective of other states that have different EMS system models, data systems, and administrative support for data collection. We used probabilistic methodology to match de-identified datasets. We may have not matched datasets optimally with this method, and would expect better results with unique identifiers. Further work to explore the details of the matching methodology (e.g., change weighting of data element; vary details of match, such as age +/− 5 years vs. one year, etc.; construction of receiver operating curves to identify the optimal cut point for match/no match determination) would also be logical future directions for this area of inquiry.

Some cardiac arrests could have occurred in the emergency department and would not have been coded by EMS as a field cardiac arrest. We anticipated that the MIDB, as a well-funded database with dedicated staff, would provide high quality, but we were chagrined to find that at least one hospital system did not submit data to it.

CONCLUSION

Probabilistic linkage of out-of-hospital cardiac arrest records and inpatient databases are feasible and enables characterization of in-hospital care of resuscitated out of hospital cardiac arrest patients. Use of the NEMSIS based EMS database and the Michigan Inpatient database was feasible and yielded a fair match compared to existing regional cardiac arrest databases. We conclude state NEMSIS datasets are imperfect to describe and evaluate statewide population based cardiac arrest care. Investment in resources to enhance EMS data collection is needed to allow their use as community measures of health. Simple (but not easy) strategies such as mandatory EMS unique identifiers are needed to link field and hospital data. This study highlights that further work is needed to overcome the practical challenges of developing integrated databases that are needed to successfully measure and improve the care of cardiac arrest patients.

REFERENCES


学霸图书馆
www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页  文献云下载  图书馆入口  外文数据库大全  疑难文献辅助工具