Taylor & Francis
Taylor & Francis Group

Prehospital Emergency Care

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ipec20

Linkage of Emergency Medical Services and Hospital Data: A Necessary Precursor to Improve Understanding of Outcomes of Prehospital Care

I. E. Blanchard, T. S. Williamson, P. Ronksley, B. Hagel, D. Niven, S. Dean, M. N. Shah, E. S. Lang & C. J. Doig

To cite this article: I. E. Blanchard, T. S. Williamson, P. Ronksley, B. Hagel, D. Niven, S. Dean, M. N. Shah, E. S. Lang & C. J. Doig (2022) Linkage of Emergency Medical Services and Hospital Data: A Necessary Precursor to Improve Understanding of Outcomes of Prehospital Care, Prehospital Emergency Care, 26:6, 801-810, DOI: 10.1080/10903127.2021.1977438

To link to this article: https://doi.org/10.1080/10903127.2021.1977438

Published online: 20 Oct 2021.

Linkage of Emergency Medical Services and Hospital Data: A Necessary Precursor to Improve Understanding of Outcomes of Prehospital Care

I. E. Blanchard, PhD, ACP, T. S. Williamson, PhD, P. Ronksley, PhD, B. Hagel, PhD, D. Niven, MD, PhD, S. Dean, PhD, M. N. Shah, MD, MPH, E. S. Lang, MD, C. J. Doig, MD, MSc

Abstract

Objective: Linking emergency medical services (EMS) data to hospital outcomes is important for quality assurance and research initiatives. However, non-linkage due to missing or incomplete patient information may increase the risk of bias and distort findings. The purpose of this study was to explore if an optimization strategy, in addition to an existing linkage process, improved the linkage rate and reduced selection and information bias. **Methods:** 4,150 transported patients in a metropolitan EMS system in Alberta, Canada from 2016/17 were linked to two Emergency Department (ED) databases by a standard strategy using a unique health care number, date/time of ED arrival, and hospital name. An optimized strategy added additional linkage steps incorporating last name, year of birth, and a manual search. The strategies were compared to assess the rate of linkage, and to describe event and patient-level characteristics of unlinked records. **Results:** The standard strategy resulted in 3,650 out of 4,150 (88.0%) linked records (95% CI 86.9%-88.9%). Of the 500 non-linked records, an additional 381 were linked by the optimized strategy (n = 4,031/4,150 [97.1%; 95% CI: 96.6%-97.6%]). There were no false positive linkages. The highest linkage failure was in 25 to 34 year-old patients (n = 93/478, 19.5%), males (n = 236/1975, 12.0%), Echo level events (n = 15/77, 19.5%), and emergency transport (45/231, 19.5%). The optimized strategy improved linkage in these groups by 68.8% (64/93), 79.2% (187/236), 40.0% (6/15), and 51.1% (23/45) respectively. For dispatch card, the highest linkage failure occurred in Card 24-Pregnancy/Childbirth/Miscarriage (n = 30/44, 68.2%), Card 27-Stab/Gunshot/Penetrating Trauma (n = 6/17, 35.3%), and Card 9-Cardiac/Respiratory Arrest/Death (n = 12/46, 26.1%). The optimized strategy improved linkage by 10.0% (3/30), 83.3% (5/6), and 41.7% (5/12) respectively. For the 119 unlinked records, 71 (59.7%) had sufficient information for linkage, but no appropriately matching records could be found. **Conclusion:** An optimized sequential deterministic strategy linking EMS data to ED outcomes improved the linkage rate without increasing the number of false positive links, and reduced the potential for bias. Even with adequate information, some records were not linked to their ED visit. This study underscores the importance of understanding how data are linked to hospital outcomes in EMS research and the potential for bias. **Key words:** emergency medical services; paramedic; ambulance; data linkage; bias

PREHOSPITAL EMERGENCY CARE 2022;26:801–810

Received April 6, 2021 from Emergency Medical Services, Emergency, Critical Care, Alberta Health Services, Alberta, Canada (IEB, DN, SD, ESL, CJD); Cumming School of Medicine - Community Health Sciences, Critical Care, Emergency, and Pediatrics; Faculty of Kinesiology, Sports Injury Prevention Research Centre, University of Calgary, Alberta Children's Hospital Research Institute and O'Brien Institute for Public Health, Calgary, Alberta, Canada (IEB, TSW, PR, BH, DN, SD, ESL, CJD); Emergency Medicine, School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin, USA (MNS). Revision received September 1, 2021; accepted for publication September 1, 2021.

Ian Blanchard was supported by a Killam Scholarship for this study.

The authors report no conflicts of interest.

IEB would like to thank the Killam family and acknowledge the Killam Pre-Doctorial Scholarship program for support that directly contributed to the completion of this study. We would like to thank Alberta Health Services, Emergency Medical Services (AHS EMS) Chief Paramedic Darren Sandbeck and Senior Medical Director Dr. Mark MacKenzie for unwavering support of paramedic science and research. Also thanks to Dr. Hude Quan, Dr. Peter Faris, Dr. Daniel Lane, and Mr. Ryan Lee for feedback and suggestions that improved this study. Mr. Greg Vogelaar, Mr. Bing Li, and Ms. Dongmei Wang provided data retrieval and assistance which was most appreciated.

Finally, we would like to acknowledge and thank the frontline paramedics and emergency communications officers who diligently collect the information upon which this study is based.

Address correspondence to Ian Blanchard, Alberta Health Services, Emergency Medical Services, 2626 Country Hills Blvd. NE, Calgary, Alberta T3N1A7, USA. E-mail: Ian.blanchard@ahs.ca

doi:10.1080/10903127.2021.1977438

Introduction

The evolving complexity of the scope of practice among paramedics, combined with the sophistication of emergency medical services (EMS) systems, requires rigor in quality assurance (QA) and research to align with and further develop best practice (1, 2). Outcome information, such as mortality and morbidity, is critical for these activities, requiring linked prehospital and hospital clinical records.

National EMS research agendas from the United States (US) and Canada identified access to patient outcome information as lacking (3–9). The US EMS Research Agenda described data linkage as critical to compile and evaluate patient outcomes (7). The Canadian EMS Research Agenda listed data linkage as both a barrier to EMS research, and as one of 19 recommendations to improve the research enterprise in EMS (3). The US National EMS Management Association (NEMSMA) published a position statement advocating for bilateral data sharing and linkage between EMS and receiving hospitals, and described multiple barriers including legal, technological, political and economic factors (10).

Record linkage connects a patient in a master database to their correct record in another database (11–13). Broadly, record linkage approaches are classified as deterministic or probabilistic (11, 12, 14). Deterministic linkage uses a series of rules that compare individual attributes (e.g., last name, age) to determine linkage. These approaches may comprise a single step (exact deterministic) strategy, or multiple step strategies (iterative deterministic), with the latter involving sequential rounds of matching (12). Probabilistic linkage uses conditional probabilities to determine the likelihood of two records being from the same individual. Probabilistic linkage also allows for certain attributes to have higher weight than others in the probability calculation (12). Unlike a deterministic approach, probabilistic linkage allows for uncertainty, and a threshold for linkage can be specified. There are no standards as to which linkage approach is most appropriate. When direct identifiers are available, and data are of reasonable quality, a deterministic approach is recommended. When direct identifiers are not available, or data quality are poor, probabilistic linkage is recommended (12). A combination of the two may also be employed (11, 12).

Linking EMS to hospital datasets must connect the correct patient and the correct encounter. Linking on patient only may generate multiple successful matches, but not all records may pertain to the correct EMS encounter. Only a few studies have described the process of linking EMS to hospital data, reporting heterogeneous cohorts, approaches and mixed results. Several Australian reports used mixed deterministic and probabilistic approaches to link all transported and non-transported patients in an EMS system with a 97.2% success rate (15, 16). US studies from Arizona and western Pennsylvania reported a 91.2% and 95.0% linkage rate respectively for all transported encounters using deterministic approaches, and a study from Oregon in trauma patients reported a sensitivity of 96.1% using a probabilistic approach (17–19). In Canada, an early record linkage study reported a 91.0% linkage rate in transported cardiac arrest patients using a probabilistic approach (20), and in the United Kingdom cardiac arrest linkage employing deterministic, probabilistic, and manual methods resulted in an 86.7% linkage rate using randomized controlled trial data, and 80.5% using registry data (21, 22).

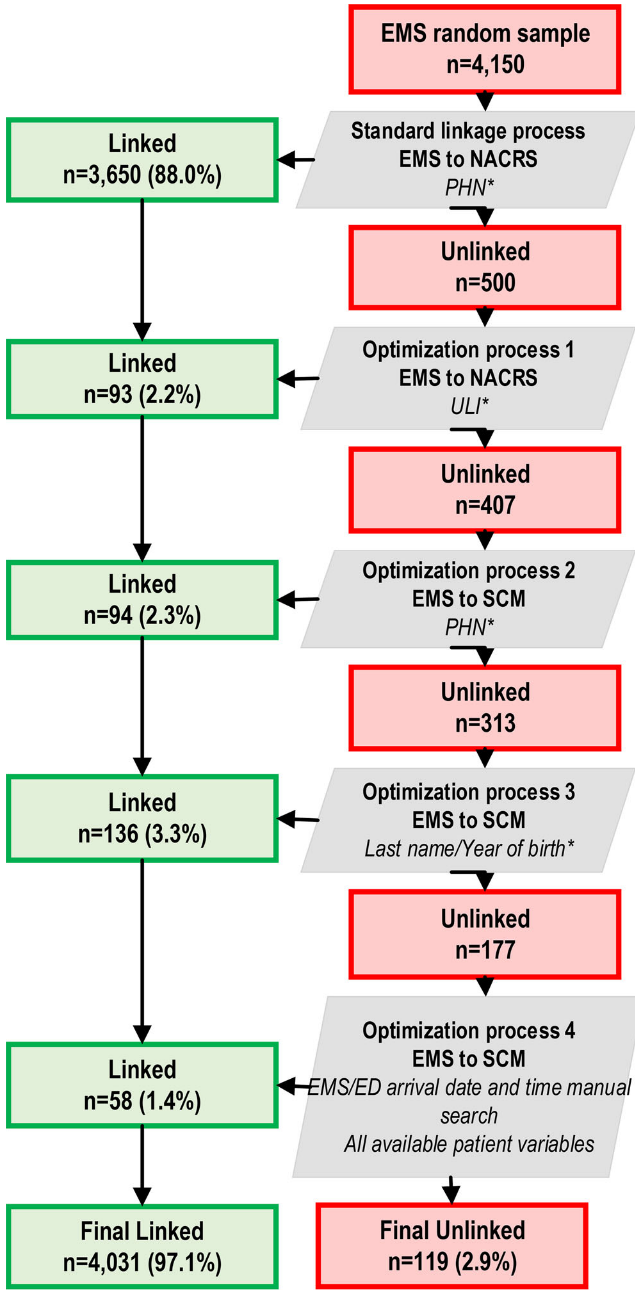
Bias from non-linkage has been raised as an issue in the administrative data literature (23, 24), EMS studies have compared linked to non-linked records noting similarities and discrepancies in event and patient characteristics (15, 19–21), and others have noted challenges in EMS linkage (25, 26). However, no known studies have compared cohorts created by different linkage strategies, or assessed if further linkage beyond a unique population level patient identifier reduces the potential for bias. Selection bias and limitations in generalizability outside of a study sample may be introduced from linkage if certain patient groups, or patients with certain characteristics are systematically not linked. For instance, critically ill patients may not have identifying information essential for record linkage recorded, therefore potentially biasing outcome assessments through selecting only those who link to be part of the analysis. Linkage strategies that obtain a high linkage rate, but with a high rate of linkage to the wrong record in certain groups, may introduce information bias, by creating erroneous information. Finally statistical power may be reduced with non-linkage, especially in the linkage of specific clinical cohorts that may be rare events.

To monitor and improve quality of care and to conduct high quality research, efficiently linking EMS data with outcomes in hospital administrative health datasets without introducing bias is important. The purpose of this study was to explore if the addition of a linkage optimization strategy improved the linkage rate and reduced the potential for selection and information bias.

METHODS

Setting and Data Systems

The Calgary Zone of Alberta Health Services (AHS) EMS serves approximately 1.5 million people in a 15,200 square mile (39,300 square kilometre) area; encompassing a mostly metropolitan setting with some suburban and rural population centers. AHS



Note: EMS=Emergency Medical Services; NACRS=National Ambulatory Care Reporting System; PHN=Provincial Health Number; SCM=Sunrise Clinical Manager; ULI=Unique Lifetime Identifier

**Also linked to EMS/Emergency Department arrival and transport destination*

Figure 1. Linkage rate by optimization step (n = 4,150).

operates 14 hospitals in this zone, including one pediatric and four adult metropolitan hospitals, and two metropolitan urgent care centers (27). AHS EMS uses the Advanced Medical Priority Dispatch (AMPDS) system, and annually coordinates the response to 160,000 emergency and non-emergency events by basic (BLS) and advanced life support (ALS) response and transfer ground vehicles (28–30). Paramedics use a single set of provincial protocols and a standardized electronic patient care record system (ePCR Suite, Version 3.8, Medusa Medical Technologies) (31).

Data were captured in two primary EMS databases. Information such as the time of event, vehicle movements, and assessment of the 9-1-1 call were recorded into the Computer Aided

Dispatch (CAD) database. Demographic and clinical information on the assessment and treatment of the patient were recorded on the ePCR and uploaded to a database. The two databases are routinely connected using a unique patient identifier by a SAP Business Objects platform (32). This unique EMS specific patient identifier is created by the CAD system at the time of the 9-1-1 call, and the connected data are used for quality, analytics, and research.

Hospital outcomes data were contained within two AHS data sources with information on ED care and disposition: the National Ambulatory Care Reporting system (NACRS) and Sunrise Clinical Manager (SCM). These datasets do not include information on hospital discharge data for inpatients, which is contained in another dataset. NACRS contains a patient's Provincial Health Number (PHN – A unique number afforded to those covered by the Alberta Health Care Insurance Plan, which is a provincial health care plan covering all insured health services in the province of Alberta, Canada) and hospital arrival time/date/hospital name, but little additional identifying information (33). SCM contains personally identifying information such as name and date of birth, but was not available for hospitals and urgent care centers outside of the metropolitan center (32).

Study Cohort

The study cohort was derived from one fiscal year (April 1, 2016 to March 31, 2017) of 9-1-1 call data (n = 116,400), restricted to transported events (n = 76,042), limited to the five metropolitan hospitals and two metropolitan urgent care centers (n = 69,769) and then randomly sampled within the remaining events (n = 4,150) using the '*runiform*' command in STATA (STATA Corp., College Station, Texas) (34). The sample size was calculated on the assumption that the standard linkage strategy would be successful for approximately 90% of records (35). An optimized linkage strategy that increased linkage by 385 patients was estimated to provide a margin of error of ±5%, assuming maximum data variability (0.50). An additional 30 patients were added, to compensate for potentially missing linkage variables. Therefore, a sample size of 415 records for optimization was anticipated to be achieved with a study cohort of 4,150.

Data Preparation and Linkage Strategies

Candidate linkage variables were identified by a comparison of the EMS and hospital databases. To link the patient to their respective hospital record there were five potential patient-level variables (PHN, name, date of birth, sex, and hospital record number), and two potential encounter-level variables (time and date of hospital arrival, and hospital name).

The standard EMS linkage strategy used in Alberta—an iterative deterministic approach using sequential rounds of matching between EMS and NACRS on PHN, EMS/ED arrival within two hours, and EMS transport destination—was performed. Records that did not match on PHN using the standard strategy, were subject to an 'optimized strategy' involving an attempt to link by PHN to another NACRS variable called the Unique Lifetime Identifier (ULI). Records that remained unlinked were then linked by PHN to SCM, followed next by last name and year of birth. Finally, for records that remained unlinked, a manual search was conducted for the shortest EMS/ED arrival difference (see Figure 1). Patients were considered linked if the PHN exactly matched, and arrival was within two hours to the correct facility. For the 'optimized strategy', the last name and year of birth had to exactly match, and arrival was within two hours to the correct facility. For the small number of cases requiring a manual search, the reviewer used all available variables to form a gestalt impression of whether the file matched, erring on the side of no match in uncertain cases.

A final manual review was conducted by one investigator (IEB) on each linked pair to confirm correct linkage to both patient and encounter using all candidate variables (i.e., looking for false matches). Although not required, a second adjudication by another team member (CJD) was possible if there was still uncertainty as to whether a correct linkage had taken place.

Analysis and Outcomes

The results from the standard linkage strategy were compared with the standard plus optimized strategy. Outcome measures included the difference between the two strategies in the proportion of linked records and false positive linked records, the proportion of linked records after each round of linking, the difference in proportion between the full cohort and linked sample characteristics, and characteristics of those who remained unlinked.

Bias from systematically low linkage in certain subgroups, was assessed using the difference between the study cohort and the samples generated by each linkage strategy for a priori defined EMS variables pertaining to conditions thought to describe critically ill or injured patients, or patients presenting with minimal personal identifiers. Event-level variables included

TABLE 1. Event-level characteristics of the study, standard linkage, and optimized linkage cohorts

Characteristic	Cohorts			Unlinked encounters		
	Study (n = 4,150) n(%)*	Standard (n = 3,650) n(%)*	Optimized (n = 4,031) n(%)*	Standard n(%)†	Optimized n(%)†	Difference n(%)†
Dispatch Priority (n = 4,149)						
Echo	77 (1.9%)	62 (1.7%)	68 (1.7%)	15 (19.5%)	9 (11.7%)	6 (7.8%)
Delta	1,112 (26.8%)	951 (26.1%)	1,063 (26.4%)	161 (14.5%)	49 (4.4%)	112 (10.1%)
Charlie	1,159 (27.9%)	1,038 (28.5%)	1,139 (28.3%)	121 (10.4%)	20 (1.7%)	101 (8.7%)
Bravo	610 (14.7%)	527 (14.4%)	592 (14.7%)	83 (13.6%)	18 (3.0%)	65 (10.7%)
Alpha	1,161 (28.0%)	1,046 (28.7%)	1,140 (28.3%)	115 (9.9%)	21 (1.8%)	94 (8.1%)
Omega	30 (0.7%)	25 (0.7%)	28 (0.7%)	5 (16.7%)	2 (6.7%)	3 (10.0%)
Dispatch Card (n = 4,150)						
1. Abdominal Pain/Problems	220 (5.3%)	190 (5.2%)	217 (5.4%)	30 (13.6 %)	3 (1.4%)	27 (12.3%)
2. Allergies/Envenomation	42 (1.0%)	39 (1.1%)	42 (1.0%)	3 (7.1%)	0 (0.0%)	3 (7.1 %)
3. Animal Bites/Attacks	3 (0.1%)	3 (0.1%)	3 (0.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4. Assault/Sexual Assault/Stun Gun	60 (1.5%)	45 (1.2%)	58 (1.4%)	15 (25.0%)	2 (3.3%)	13 (21.7%)
5. Back Pain (Non-Traumatic)	65 (1.6%)	58 (1.6%)	65 (1.6%)	7 (10.8 %)	0 (0.0%)	7 (10.8%)
6. Breathing Problems	362 (8.7%)	325 (8.9%)	358 (8.9%)	37 (10.2%)	4 (1.1 %)	33 (9.1%)
7. Burns/Explosion	6 (0.1%)	6 (0.2%)	6 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0 %)
8. Carbon Monoxide/HAZMAT	4 (0.1%)	4 (0.1%)	4 (0.1%)	0 (0.0 %)	0 (0.0 %)	0 (0.0%)
9. Cardiac/Respiratory Arrest/Death	46 (1.1%)	34 (0.9%)	39 (1.0%)	12 (26.1%)	7 (15.2 %)	5 (10.9%)
10. Chest Pain/Chest Discomfort	425 (10.2%)	384 (10.5%)	417 (10.3%)	41 (9.6%)	8 (1.9%)	33 (7.8%)
11. Choking	18 (0.4%)	18 (0.5%)	18 (0.5%)	0 (0.0%)	0 (0.0 %)	0 (0.0%)
12. Convulsions/Seizures	147 (3.5%)	126 (3.5%)	142 (3.5%)	21 (14.3%)	5 (1.4 %)	16 (10.9%)
13. Diabetic Problems	42 (1.0%)	39 (1.1%)	41 (1.0%)	3 (7.1%)	1 (2.4%)	2 (4.8%)
14. Drowning/SCUBA Accident	1 (<0.1%)	1 (<0.1%)	1 (<0.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
16. Eye Problems/Injuries	6 (0.1%)	5 (0.1%)	6 (0.2%)	1 (16.7%)	0 (0.0%)	1 (16.7%)
17. Falls	358 (8.6%)	327 (9.0%)	353 (8.8%)	31 (8.7%)	5 (1.4%)	26 (7.3%)
18. Headache	27 (0.7%)	25 (0.7%)	26 (0.7%)	2 (7.4%)	1 (3.7%)	1 (3.7%)
19. Heart Problems/A.I.C.D.	70 (1.7%)	67 (1.8%)	70 (1.7%)	3 (4.3%)	0 (0.0%)	3 (4.3%)
20. Heat/Cold Exposure	5 (0.1%)	5 (0.1%)	5 (0.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
21. Hemorrhage/Lacerations	177 (4.3%)	156 (4.3%)	175 (4.3%)	21 (11.9%)	2 (1.1%)	19 (10.7%)
23. Overdose/Poisoning	133 (3.2%)	111 (3.0%)	129 (3.2%)	22 (16.5%)	4 (3.0%)	18 (13.5%)
24. Pregnancy/Childbirth	44 (1.1%)	14 (0.4%)	17 (0.4%)	30 (68.2%)	27 (61.4%)	3 (6.8%)
25. Psychiatric/Suicide	141 (3.4%)	113 (3.1%)	136 (3.4%)	28 (19.9%)	5 (3.5%)	23 (16.3%)
26. Sick Person	627 (15.1%)	564 (15.5%)	616 (15.3%)	63 (10.0%)	11 (1.8%)	52 (8.3%)
27. Stab/Gunshot	17 (0.4%)	11 (0.3%)	16 (0.4%)	6 (35.3%)	1 (5.9%)	5 (29.4%)
28. Stroke/TIA	110 (2.7%)	104 (2.9%)	109 (2.7%)	6 (5.5%)	1 (0.9%)	5 (4.5%)
29. Traffic/Transportation Incidents	123 (3.0%)	108 (3.0%)	118 (2.9%)	15 (12.2%)	5 (4.1%)	10 (8.1%)
30. Traumatic Injuries	134 (3.2%)	127 (3.5%)	132 (3.3%)	7 (5.2%)	2 (1.5%)	5 (3.7%)
31. Unconscious/Fainting	292 (7.1%)	249 (6.8%)	280 (7.0%)	43 (14.7%)	12 (4.1%)	31 (10.6%)
32. Unknown Problem	50 (1.2%)	40 (1.1%)	46 (1.1%)	10 (20.0%)	4 (8.0%)	6 (12.0%)
33. Transfer/Palliative Care	395 (9.5%)	352 (9.6%)	386 (9.6%)	43 (10.9%)	9 (2.3%)	34 (8.6%)
Hospital Transport Priority (n = 4,118)						
Lights and Siren	231 (5.6%)	186 (5.1%)	209 (5.2%)	45 (19.5%)	22 (9.5%)	23 (10.0%)
Non-emergency	3,887 (94.4%)	3,454 (94.9%)	3,811 (94.8%)	433 (11.1%)	76 (2.0%)	357 (9.2%)

Note: AICD = automatic implantable cardioverter/defibrillator; EtOH = Ethyl alcohol; HAZMAT = hazardous materials; LOC = level of consciousness; MVC = motor vehicle collision; TIA = transient ischemic attack.
*Within column percentage.
†Between column percentage with the study cohort as denominator.

dispatch determinant (the priority at which EMS responded), dispatch card (the primary reason for the 9-1-1 call as interpreted by the Emergency Medical Dispatcher), and hospital transport priority (lights and siren or non-emergency). Patient-level variables included age, sex, and patient chief complaint (as described by the patient/event characteristics and recorded by the paramedic).

All analyses were performed using STATA Version 15.1 (STATA corp., College Station, Texas). All statistical tests were considered significant at an alpha level of 0.05.

Ethical Considerations

This research was approved and overseen by the University of Calgary Conjoint Health Research Ethics Board (Ethics Certificate number 17-1929) and met the requirements for waiver of consent. This study utilized personally identifying health

TABLE 2. Patient-level characteristics of the study, standard linkage, and optimized linkage cohorts

	Cohorts			Unlinked encounters		
Characteristic	Study (n = 4,150) n(%)*	Standard (n = 3,650) n(%)*	Optimized (n = 4,031) n(%)*	Standard n(%)†	Optimized n(%)†	Difference n(%)†
Age (n = 4,114)						
0 to 2 years	80 (1.9%)	74 (2.0%)	78 (1.9%)	6 (7.5%)	2 (2.5%)	4 (5.0%)
3 to 17 years	211 (5.1%)	190 (5.2%)	205 (5.1%)	21 (10.0%)	6 (1.5%)	15 (7.1%)
18 to 24 years	265 (6.4%)	228 (6.3%)	261 (6.5%)	37 (14.0%)	4 (2.8%)	33 (12.5%)
25 to 34 years	478 (11.6%)	385 (10.6%)	449 (11.1%)	93 (19.5%)	29 (6.1%)	64 (13.4%)
35 to 64 years	1,350 (32.8%)	1,191 (32.6%)	1,319 (32.7%)	159 (11.8%)	31 (2.3%)	128 (9.5%)
65 years and older	1,730 (42.1%)	1,582 (43.3%)	1,718 (42.6%)	148 (8.6%)	12 (0.7%)	136 (7.9%)
Sex (n = 4,124)						
Male	1,975 (47.9%)	1,739 (47.6%)	1,926 (47.8%)	236 (11.9%)	49 (2.5%)	187 (9.5%)
Female	2,149 (52.1%)	1,911 (52.4%)	2,105 (52.2%)	238 (11.1%)	44 (2.0%)	194 (9.0%)
Chief Complaint‡ (n = 4,105)						
Altered LOC	125 (3.1%)	108 (3.0%)	120 (3.0%)	17 (13.6%)	5 (4.0%)	12 (9.6%)
Bleeding	63 (1.5%)	58 (1.6%)	62 (1.5%)	5 (7.9%)	1 (1.6%)	4 (6.3%)
Blood Sugar – Low	19 (0.5%)	18 (0.5%)	19 (0.5%)	1 (5.3%)	0 (0.0%)	1 (5.3%)
Burn	7 (0.2%)	7 (0.2%)	7 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Cardiac Arrest	14 (0.3%)	11 (0.3%)	13 (0.3%)	3 (21.4%)	1 (7.1%)	2 (14.3%)
Choking	19 (0.5%)	19 (0.5%)	19 (0.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Drowning/Near Drowning	2 (0.1%)	2 (0.1%)	2 (0.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
MVC	83 (2.0%)	76 (2.1%)	81 (2.0%)	7 (8.4%)	2 (2.4%)	5 (6.0%)
Poisoning/Overdose	143 (3.5%)	119 (3.3%)	138 (3.4%)	24 (16.8%)	5 (3.5%)	19 (13.3%)
Pregnancy/Childbirth	18 (0.4%)	1 (<0.1%)	2 (0.1%)	17 (94.4%)	16 (88.9%)	1 (5.6%)
Psychiatric Problems	206 (5.0%)	173 (4.8%)	198 (4.9%)	33 (16.0%)	8 (3.9%)	25 (12.1%)
Respiratory Arrest	1 (<0.1%)	1 (<0.1%)	1 (<0.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Seizures	121 (3.0%)	105 (2.9%)	120 (3.0%)	16 (13.2%)	1 (0.8%)	15 (12.4%)
Shortness of Breath	258 (6.3%)	238 (6.5%)	258 (6.4%)	20 (7.8%)	0 (0.0%)	20 (7.8%)
Stroke/TIA	52 (1.3%)	44 (1.2%)	51 (1.3%)	8 (15.4%)	1 (1.9%)	7 (13.5%)
Substances Abuse – EtOH	116 (2.8%)	96 (2.6%)	108 (2.7%)	20 (17.2%)	8 (6.9%)	12 (10.3%)
Syncope/Near Syncope	135 (3.3%)	119 (3.3%)	132 (3.3%)	16 (11.9%)	3 (2.2%)	13 (9.6%)
Trauma	331 (8.1%)	291 (8.0%)	327 (8.2%)	40 (12.1 %)	4 (1.2%)	36 (10.9%)
Unresponsive	17 (0.4%)	14 (0.4%)	15 (0.4%)	3 (17.6%)	2 (11.8%)	1 (5.9%)

Note: AICD = automatic implantable cardioverter/defibrillator; EtOH = Ethyl alcohol; HAZMAT = hazardous materials; LOC = level of consciousness; MVC = motor vehicle collision; TIA = transient ischemic attack.
*Within column percentage.
†Between column percentage with the study cohort as denominator.
‡Chief Complaint as reported by the patient/event characteristics and recorded by the paramedic.

information, and therefore the following procedures were put in place to mitigate the risk of unintended breach of confidentiality: identifying datasets were kept behind the firewall of the health care system, no health information were attached to the identifying information, the minimum number of investigators had access to identifying information, and privacy agreements were in place.

RESULTS

Study Cohort Characteristics

The event-level and patient-level characteristics of the study cohort are outlined in Tables 1 and 2 respectively. Briefly, approximately 30% of events were for situations thought to be immediately life-threatening (Echo or Delta determinants) or minor (Alpha or Omega determinants) respectively. The most common dispatch cards were sick person (15.1%), chest pain/discomfort (10.2%), and transfer/palliative care (9.5%). Almost 95% of events resulted in non-emergency transport. The largest age group were those over the age of 65 years (42.1%), with slightly more females (52.1%). The most common chief complaints were trauma (8.1%), shortness of breath (6.3%), and psychiatric problems (5.0%).

Linkage Strategy Samples

Figure 1 outlines the proportion of records that were linked. The standard linkage strategy resulted in 3,650 out of 4,150 (88.0%) linked records (95% CI 86.9%, 88.9%). Of the 500 non-linked records, an additional 381 were linked after the optimized strategy, or 4,031 out of 4,150 (97.1%) linked records

TABLE 3. Missing linkage variables and characteristics of records that remained unlinked with the optimized linkage strategy (n = 119)

Variable*	n (%)
PHN	58 (48.7%)
Last name	48 (40.3%)
First name	44 (37.0%)
Date of Birth	51 (42.9%)
Hospital Record Number	101 (84.9%)
EMS/ED arrival	0 (0.0%)
EMS Transport destination	0 (0.0%)
Characteristic	**n (%)**
Age (n = 84)	
0 to 2 years	2 (2.4%)
3 to 17 years	6 (7.1%)
18 to 24 years	4 (4.8%)
25 to 34 years	29 (34.5%)
35 to 64 years	31 (36.9%)
65 years and older	12 (14.3%)
Sex (n = 93)	
Male	49 (52.7%)
Female	44 (47.3%)
Dispatch Priority (n = 119)	
Echo	9 (7.6%)
Delta	49 (41.2%)
Charlie	20 (16.8%)
Bravo	18 (15.1%)
Alpha	21 (17.7%)
Omega	2 (1.7%)
Transport Priority (n = 98)	
Emergency	22 (22.5%)
Non-emergency	76 (77.6%)

NOTE: ED = Emergency Department; EMS = Emergency Medical Services; PHN = Provincial Health Number.
*Records may contain multiple missing variables.

(95% CI 96.6%, 97.6%). Upon manual review there were no instances of false positive linkages in either strategy.

Table 1 outlines the event-level characteristics of the sample resulting from the standard and optimized linkage strategies. The highest proportion of linkage failure with the standard strategy occurred in Echo level events (n = 15/77, 19.5%), with an additional six out of the 15 linked by the optimized strategy. For dispatch card, the three cards with the highest proportion of linkage failure with the standard strategy occurred in Card 24 – Pregnancy/Childbirth/Miscarriage (n = 30/44, 68.2%), Card 27 – Stab/Gunshot/Penetrating Trauma (n = 6/17, 35.3%), and Card 9 – Cardiac or Respiratory Arrest/Death (n = 12/46, 26.1%). The optimized strategy linked an additional three out of 30, five out of six, and five out of 12 respectively. For transport priority more linkage failure occurred in those receiving lights and siren hospital transport (45/231, 19.5%) compared to non-emergency. The optimized strategy linked an additional 23 out of 45.

Table 2 outlines patient-level characteristics, with the highest proportion of linkage failure occurring in 25 to 34 year olds (n = 93/478, 19.5%), and males (n = 236/1975, 12.0%). The optimized strategy linked an additional 64 out of 93, and 187 out of 236 respectively. The top three Chief Complaints with the highest proportion of linkage failure with the standard strategy were Pregnancy/Childbirth (n = 17/18, 94.4%), Cardiac Arrest (n = 3/14, 21.4%), and Unresponsive (n = 3/17, 17.7%). The optimized strategy linked an additional one out of 17, two out of three, and one out of three respectively.

Unlinked Records

There were 119 out of 4,150 (2.9%, 95% CI 2.4%, 3.4%) records not linked after the optimized strategy. In 71 (59.7%) of these records, there was sufficient information present for matching, but no appropriately matching records could be found in either NACRS or SCM. The remaining 48 records (40.3%) were exclusively missing patient information, but had encounter information present. Table 3 describes missing data and EMS recorded characteristics of the unlinked records. The top missing linkage variable was Hospital Record Number, which even when present could not be readily used because of multiple different formats (i.e., paramedics recorded variations of the number compared to hospital datasets) and unrelated information (i.e., paramedics used this field as free text for other information). Of the 58 records that were missing PHN, 48 were also missing last name. Of the 10 records that were not missing last name, seven also had first name and date of birth but could still not be matched.

DISCUSSION

Use of the optimized linkage strategy in addition to the standard strategy, increased the linkage rate from 88.0% to 97.1% (9.1%) without increasing false positive links. There was evidence that the standard plus optimized strategy increased linkage in important event and patient-level characteristic subgroups such as those who are thought to be in a life-threatening situation at the time of the 9-1-1 call, trauma, cardiac arrest, and those transported using lights and siren. These results compare favorably with published literature, which reports a range of linkage rates from approximately 80% to 97% (15–22). This range is likely influenced by factors such as the variables available for linkage (e.g., a unique patient identifier between EMS and the hospital, etc.), quality of the data, the linkage strategy, and inclusion criteria (e.g., all EMS patients versus a clinical cohort like cardiac arrest, transported patients only versus all patients, etc.).

The issue of bias from unlinked records has been raised in the administrative data literature in other topic and speciality areas (24). A narrative review by Bohensky and colleagues in 2010 found sex, race, geographical/hospital site, socio-economic status, and health status differed between linked and unlinked records (23). In EMS, studies have reported that sex, age, service area (i.e., metropolitan or regional), location (i.e., home or public place), witnessed status (i.e., not witnessed, bystander, EMS), EMS response time, transport location, pension status, acuity, clinical outcome (i.e., return of spontaneous circulation), dispatch card (i.e., other/unknown, trauma, mental health, drugs and alcohol), and paramedic protocol (i.e., drug or toxicology) may differ between linked and unlinked records (15, 19–21). Differential linkage of records over important event and patient characteristics may bias prevalence estimates and estimates of association, with the direction and magnitude dependent on the research question under consideration and extent of non-linkage (23, 24). This underscores the necessity of a robust linkage process to mitigate this theoretical risk.

In the EMS setting, our study suggests that the potential for selection bias could exist for certain subgroups. For example, of the 44 patients that presented with a Card 24 – Pregnancy/Childbirth, 30 (68.2%) could not be linked. This was only reduced to 27 (61.4%) unlinked with the optimized strategy. A potential explanation for this low linkage rate was a destination protocol in our system for pregnant patients between 20 to 32 weeks gestation that leads paramedics to transport directly to the Labor and Delivery Unit at one hospital (31). This "ED bypass" would logically exclude these women from ED datasets, and systematically exclude them from studies that used this linkage strategy. Detailed understanding of local clinical processes that might affect data linkage is important, and additional linkage of EMS data directly to in-patient datasets in addition to ED datasets may mitigate this impact. Other subgroups that had low linkage, but were more amenable to improved linkage using the optimized strategy included trauma and cardiac arrest patients. This was not surprising given that it is often difficult to ascertain detailed information on-scene in these circumstances as paramedic attention is focused on rapid assessment and care, rather than confirming a PHN for example.

EMS linkage must link the right patient to their record, and also the right encounter. False-matches, where a record is linked that should not have been linked, may also distort prevalence estimates and associations from information bias (24). Like selection bias, the extent of information bias may be dependent on the linkage strategy, study inclusion criteria, and exposure and outcome under consideration. For this study, manual review identified no false-positive matches. This was not surprising as compared to probabilistic approaches that are often employed on less robust linkage data, deterministic linkages may not cause an increased risk of information bias (18).

Of the records that could not be linked the two major issues were missing data, and a failure to locate an ED record. The former issue highlights the importance of adequate data collection and the dual and opposing roles placed on paramedics in providing timely critical care with rapid transport, while attempting to gather patient information. The latter issue underscores the importance of data accuracy (e.g., incorrectly recorded PHN, name, date of birth, etc.). Further research should explore the reasons why records could not be found in the ED dataset, in spite of patient and encounter related data from EMS. Both issues draw attention to the potential effectiveness of an automatically generated unique tracking number accurately recorded in both the EMS and hospital record on ED arrival. Such a process would obviate complex and resource intensive linkage strategies, as it would deterministically link both patient and encounter. As has been well described in research agendas (3–9), visioning documents (36–39), and position statements (10), understanding EMS patient outcomes are important. Linkage of EMS data to hospital outcomes will drive research, quality and analytics, which is essential for the development of a solid foundation of knowledge from which EMS systems, the paramedic profession, and EMS physician subspecialty can be positioned.

A limitation of this study is the difficulty in generalizing the linkage steps from a small to large sample. The small size of the sample allowed for manual review as the final optimization step. While this step increased the linkage by 58 patients or 1.4%, the issue is the time that this step takes to complete. Each unlinked record was manually searched in the SCM dataset, which would not be pragmatic from a resource perspective to implement with larger datasets. This study only included metropolitan hospitals, therefore interpretive caution is required when generalizing outside of the metropolitan area. This study assessed the linkage of EMS to ED databases, but not the linkage of ED to in-patient databases, which in some jurisdictions may require additional linkage. Important outcomes such as survival to hospital discharge and length of hospital stay are contained within these in-patient datasets. Finally, this study may have limited direct generalizability outside of systems that do not have population level health integration and tracking. However, regardless of the approach used to link data, this study highlights the significance of understanding who does not link in the context of the research question under consideration.

CONCLUSIONS

An optimized sequential deterministic strategy for linking EMS data to hospital outcome improved the linkage rate without increasing the number of false positive links and reduced the potential for bias. Despite adequate information, some records were not linked to their ED visit. Starting with linkage from EMS to the ED, this study underscores the importance of understanding how data are linked to hospital outcomes in EMS research and the assessment of the potential for bias in studies using linked datasets, especially in those with modest to low linkage rates.

ORCID

I. E. Blanchard http://orcid.org/0000-0002-3619-3196

References

1. Callaham M. Quantifying the scanty science of prehospital emergency care. Ann Emerg Med. 1997;30(6):785–90. doi:10.1016/S0196-0644(97)70049-0.
2. Shah MN. The formation of the emergency medical services system. Am J Public Health. 2006;96(3):414–23. doi:10.2105/AJPH.2004.048793.
3. Jensen JL, Bigham BL, Blanchard IE, Dainty KN, Socha D, Carter A, et al. The Canadian National EMS Research Agenda: a mixed methods consensus study. CJEM. 2013;15(2):73–82. doi:10.2310/8000.2013.130894.
4. Jensen JL, Blanchard IE, Bigham BL, Carter A, Brown R, Socha D, et al. The Canadian National EMS Research Agenda: Impact and Feasibility of Implementation of Previously Generated Recommendations. CJEM. 2015;17(5):484–90. doi:10.1017/cem.2015.29.
5. Sayre MR, White LJ, Brown LH, McHenry SD. Implementation Symposium Participants National EMS research agenda: proceedings of the implementation symposium. Acad Emerg Med. 2003;10(10):1100–8. doi:10.1197/S1069-6563(03)00362-2.
6. Sayre MR, White LJ, Brown LH, McHenry SD, National EMS Agenda Writing Team. National EMS Research Agenda. Prehosp Emerg Care. 2002;6(3 Suppl):S1–S43. (3 Suppl):
7. Sayre MR, White LJ, Brown LH, McHenry SD, National EMS Research Agenda Writing Team. The National EMS Research Agenda executive summary. Emergency Medical Services. Ann Emerg Med. 2002;40(6):636–43. National EMSRAWT. The doi:10.1067/mem.2002.129241.
8. Dainty KN, Jensen JL, Bigham BL, Blanchard IE, Brown LH, Carter AJE, Socha D, Morrison LJ. Developing a Canadian emergency medical services research agenda: a baseline study of stakeholder opinions. CJEM. 2013;15(2):83–9. doi:10.2310/8000.2013.131022.
9. Jensen JL, Blanchard IE, Bigham BL, Dainty KN, Socha D, Carter A, et al. Methodology for the development of a Canadian national EMS research agenda. BMC Emerg Med. 2011;11:15. doi:10.1186/1471-227X-11-15.
10. Gunderson MR, Florin A, Price M, Reed J. NEMSMA position statement and white paper: process and outcomes data sharing between EMS and receiving hospitals. Prehosp Emerg Care. 2021;25(2):307–313.
11. Herzog T, Scheuren F, Winkler W. Data quality and record linkage techniques. New York (NY): Springer; 2007.
12. Dusetzina S, Tyree S, Meyer A, Meyer A, Green L, Carpenter W. Linking data for health services research: A framework and instructional guide. Rockville (MD): University of North Carolina at Chapel Hill; 2014.
13. Benchimol EI, Langan S, Guttmann A, Committee RS, RECORD Steering Committee. Call to RECORD: the need for complete reporting of research using routinely collected health data. J Clin Epidemiol. 2013;66(7):703–5. doi:10.1016/j.jclinepi.2012.09.006.
14. Randall S. Enabling health research using administrative data: methodological improvements. Australia: Curtin University; 2017.
15. Carroll T, Muecke S, Simpson J, Irvine K, Jenkins A. Quantification of NSW Ambulance Record Linkages with Multiple External Datasets. Prehosp Emerg Care. 2015;19(4):504–15. doi:10.3109/10903127.2015.1025154.
16. Cone DC, Irvine KA, Middleton PM. The methodology of the Australian Prehospital Outcomes Study of Longitudinal Epidemiology (APOStLE) project. Prehosp Emerg Care. 2012;16(4):505–12. doi:10.3109/10903127.2012.689929.
17. Chikani V, Blust R, Vossbrink A, Wightman P, Bissell S, Graw J, Martinez R, Fisher B. Improving the continuum of care by bridging the gap between prehospital and hospital discharge data through stepwise deterministic linkage. Prehosp Emerg Care. 2020;24(1):1–7. doi:10.1080/10903127.2019.1604925.
18. Newgard CD. Validation of probabilistic linkage to match de-identified ambulance records to a state trauma registry. Acad Emerg Med. 2006;13(1):69–75. doi:10.1197/j.aem.2005.07.029.
19. Seymour CW, Kahn JM, Martin-Gill C, Callaway CW, Angus DC, Yealy DM. Creating an infrastructure for comparative effectiveness research in emergency medical services. Acad Emerg Med. 2014;21(5):599–607. doi:10.1111/acem.12370.
20. Waien SA. Linking large administrative databases: a method for conducting emergency medical services cohort studies using existing data. Acad Emerg Med. 1997;4(11):1087–95. doi:10.1111/j.1553-2712.1997.tb03684.x.
21. Ji C, Quinn T, Gavalova L, Lall R, Scomparin C, Horton J, et al. Feasibility of data linkage in the PARAMEDIC trial: a cluster randomised trial of mechanical chest compression in out-of-hospital cardiac arrest. BMJ Open. 2018;8(7):e021519.
22. Rajagopal S, Booth SJ, Brown TP, Ji C, Hawkes C, Siriwardena AN, et al. 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. BMJ Open. 2017;7(11):e017784. doi:10.1136/bmjopen-2017-017784.
23. Bohensky MA, Jolley D, Sundararajan V, Evans S, Pilcher DV, Scott I, Brand CA. Data linkage: a powerful research

tool with potential problems. BMC Health Serv Res. 2010;10: 346. doi:10.1186/1472-6963-10-346.
24. Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, Goldstein H. Challenges in administrative data linkage for research. Big Data Soc. 2017;4(2):2053951717745678. doi: 10.1177/2053951717745678.
25. Swor R, Qu L, Putman K, Sawyer KN, Domeier R, Fowler J, Fales W. Challenges of using probabilistic linkage methodology to characterize post-cardiac arrest care in Michigan. Prehosp Emerg Care. 2018;22(2):208–13. doi:10.1080/10903127.2017.1362086.
26. Zegre-Hemsey JK, Asafu-Adjei J, Fernandez A, Brice J. Characteristics of prehospital electrocardiogram use in North Carolina using a novel linkage of emergency medical services and emergency department data. Prehosp Emerg Care. 2019;23(6):772–9. doi:10.1080/10903127.2019.1597230.
27. Calgary Zone Stats Alberta, Canada. Calgary Zone Stats Alberta, Canada: Alberta Health Services; 2021 [cited 2021 July 24]. Available from: https://www.albertahealthservices.ca/assets/Zone/ahs-zn-calgary-map-brochure.pdf.
28. About the College Alberta, Canada. About the College Alberta, Canada: Alberta College of Paramedics; 2017. [cited 2018 October 10]. Available from: https://www.collegeofparamedics.org/the-college/.
29. Emergency Health Services Alberta, Canada. Emergency Health Services Alberta, Canada: Alberta Health; 2018. [cited 2018 October 10]. Available from: http://www.health.alberta.ca/services/EHS.html.
30. Emergency Medical Services Alberta, Canada. Emergency Medical Services Alberta, Canada: Alberta Health Services Emergency Medical Services; 2018. [cited 2018 October 10]. Available from: https://www.albertahealthservices.ca/ems/ems.aspx.
31. Alberta Health Services, Emergency Medical Services. Medical Control Protocols. https://apps.apple.com/ca/app/ahs-ems/id413312484. Alberta, Canada; 2016.
32. EDW (AHSDRR) Metadata. Alberta, Canada: Alberta Health Services; 2018. [cited 2018 Ocotber 10]. Available from: https://insite.albertahealthservices.ca/dimr/Page4773.aspx.
33. Canadian Institute for Health Information. NACRS abstracting manual, 2014–2015 edition. Ottawa, Ontairio: Canadian Institute for Health Information; 2015.
34. Runiform - uniform and nonuniform pseudorandom variates. https://www.stata.com/manuals13/m-5runiform.pdf. College Station, Texas: STATA.
35. Blanchard IE, Doig CJ, Hagel BE, Anton AR, Zygun DA, Kortbeek JB, et al. Emergency medical services response time and mortality in an urban setting. Prehosp Emerg Care. 2012;16(1):142–51. doi:10.3109/10903127.2011.614046.
36. EMS Trend Report 2019: How Will EMS Advance at Current Pace of Change?: EMS1.com and Fitch and Associates; 2019.
37. EMS Agenda 2050: A People-Centered Vision for the Future of Emergency Medical Services. Downloaded from: https://www.ems.gov/pdf/EMS-Agenda-2050.pdf. Washington, DC, USA: National Highway Traffic Safety Administration (NHTSA); 2019.
38. Taking Healthcare to the Patient: Transforming NHS Ambulance Services. Downloaded from: https://www.researchgate.net/publication/329949112_Taking_Healthcare_to_the_Patient_Transforming_NHS_Ambulance_Services. London, UK: Department of Health; 2005.
39. Public Health England and Association of Ambulance Chief Executives. Developing a Public Health Approach within the Ambulance Sector: Discussion Paper. England: Association of Ambulance Chief Executives; 2021.