CEMSIS and Trauma Data Linkage Initiative: Emergency Medical Services Authority Preventive Health and Health Services Block Grant (PHHSBG)

Emergency Medical Services Authority

California Health and Human Services Agency

June 2022

Elizabeth Basnett, Acting Director

EMSA Publication #SYS 100-06

ELIZABETH BASNETT
ACTING DIRECTOR

LOUIS BRUHNKE
CHIEF DEPUTY DIRECTOR

TOM McGINNIS
CHIEF, EMS SYSTEMS DIVISION

ADRIENNE KIM
DATA AND QUALITY IMPROVEMENT MANAGER, EMS SYSTEMS DIVISION

TORY LUPINETTI
RESEARCH DATA SPECIALIST, EMS SYSTEMS DIVISION

Table of Contents

CEMSIS AND TRAUMA DATA LINKAGE INITIATIVE

INTRODUCTION

Linking data collected by EMS personnel during patient encounters to the hospital data generated for those same patients during their subsequent in-hospital care is important in understanding whole-patient care and for informing on-going prehospital quality assurance and quality improvement measures. Moreover, there is an opportunity for expanding knowledge about traumatic injury in California because the California EMS Authority (EMSA) has access to designated trauma hospital's patient registries. Under the provisions of the Preventive Health and Health Services Block Grant (PHHSBG), EMSA is attempting to link robust prehospital and hospital data to create a path for future research and policy-making decisions. The goal of the CEMSIS and Trauma Data Linkage Initiative is to maximize the number of matched records between two databases. **The objective of this brief is to:**

1. Identify a sample population where each patient has an EMS record linked to a hospital record and highlight those shared data elements that are available for analyses.
2. Demonstrate one strategy for linking data across two separate healthcare databases as a foundation for future research.
3. Describe resulting matched records and how they could potentially be analyzed and displayed in charts and figures in future reports.

BACKGROUND

The California EMS Information System (CEMSIS) was developed by EMSA with the assistance of EMS systems partners implemented a method to capture a complete EMS event, but it remains a system that mimics NEMSIS. Local EMS agencies (LEMSAs) submit their EMS provider's emergency call data to Inland Counties Emergency Medical Agencies (ICEMA). ICEMA is contracted with EMSA to serve as the facilitator of this data via ImageTrend®. Concurrently, California implemented the National Emergency Medical Services Information System (NEMSIS) data elements and standards issued by the National Highway and Traffic Safety Administration. The most recent version, NEMSIS version 3.4.0, is a data dictionary composed of 585 data elements and is designed to be a nationwide standard for reporting EMS data. Providers must comply with CEMSIS and NEMSIS data element requirements as mandated by Health and Safety Code §1797.227.

Trauma registry data is collected by the American College of Surgeons in a nationally recognized and standardized format known as the National Trauma Data Bank (NTDB). Based on National Trauma Data Standards (NTDS) the data dictionary provides the operational definitions for the NTDB. It is useful for nationwide benchmarks and trends, evaluating hospital trauma systems and patient outcomes, and providing other valuable information regarding healthcare services. California EMSA utilizes the NTDS data dictionary and the standard inclusion criteria to identify trauma patients.

Between January 1, 2019 and June 30, 2019, there were 22,266 trauma-related incidents recorded in the CEMSIS database. Accurate linkage of EMS and other patient-centered healthcare records is essential to analyze illness and injury rates in a population. Data linkage will enable further insight into patient health outcomes, expense and resource utilization, and efficacy of prehospital interventions. This report seeks to determine how reliably EMS and hospital trauma data can be linked. They are unique databases with different criteria, especially the trauma patient registry which defines a patient that had a traumatic injury. Traumatic injury data is a subset of EMS data and further research into how these data and databases relate is warranted for data matching.

Data that was collected represent a period (January-June 2019) from a Level I trauma hospital in Riverside County, Riverside Community Hospital. Data was extracted from CEMSIS and patient registry, which are both managed by ICEMA via ImageTrend®.

The following information is an overview of what the successfully linked EMS and trauma data looks like and what information can be gained from the results. The main challenge to completing the objectives of this study was the absence and inaccessibility of unique identifiers across two specific databases that were utilized (ImageTrend® Elite and patient registry). There were also missing, duplicate, or inaccurate data entered in these databases posing significant challenges to probabilistic data matching efforts. EMSA used deterministic record linkage methods for this study.

METHODOLOGY

The methodology EMSA used to link EMS and trauma data is based on a literature review of multiple other largely successful attempts at linking patient records utilizing deterministic and probabilistic matching tactics with similar EMS and hospital databases. Common across multiple reports was the use of unique identifiers such as full names, dates of birth, social security numbers, and incident dates. A 2021 study from Blanchard, Williamson, Ronksley, Hagel, Niven, Dean, Shah, Lang, and Doig demonstrated how unique identifiers can successfully link patient-level data across multiple EMS and emergency department databases. They used "date and time of arrival" at the hospital as well as a unique healthcare number which yielded an 88% linkage rate.[i] They added an additional linkage step that included patient's name, date of birth and a manual search for potential matches that led to a 97% match rate overall.

Another study from Chikani, Blust, Vossbrink, Wightman, Bissell, Graw, and Martinez (2020) used a 17-step deterministic linkage process. The first step used first and last name, date of birth, gender, and incident date which yielded a 64% match rate.[ii] Subsequent steps resulted in a 91% total match rate which used a mixture of various patient identifiers and a Soundex function (phonetically comparative names). Lastly, a report from the University of Pittsburg Medical Center (2014) successfully linked more than 95% of EMS and hospital records within its nine-hospital healthcare system to 33 local EMS agency records.[iii]

Due to the nature of CEMSIS and subsequent accessibility to EMS and trauma data, EMSA only used a one-step deterministic approach. EMS and patient registry data only operate on NEMSIS and NTDS data elements that promote patient confidentiality, so EMSA does not have access to unique patient identifiers like full names, addresses, and social security numbers. Thus, date of birth and incident date were the only two viable data variables available among the two databases. Below is a summary of methods EMSA used to collect and analyze the linked EMS and trauma registry data for Riverside Community Hospital for January 1, 2019 through June 30, 2019.

Phase One

The reports for this project were from CEMSIS using NEMSIS data standards v3.4.0 and the ImageTrend® Patient Registry using the most recent version of the NTDS data standards. EMSA did not filter data variables, so the largest sample population was obtained. For example, EMSA did not filter out PCRs that indicated "non-emergent" or "911 response" only because these patients were still queried in the trauma registry. Furthermore, a CEMSIS EMS record could have left out trauma-related information but was included in the patient's trauma record at the hospital.

EMSA staff sought a unique identifier that would yield a high match rate for EMS and trauma records. Ideally, criteria like a unique alpha-numeric sequence, first and last name, date of birth, or social security number are the most helpful. However, the only viable and accessible variables were the patient's date of birth and incident date across these two databases. EMSA was unsuccessful in linking the **"ePCR (eResponse.03)"** data element in CEMSIS to the corresponding hospital record because the **"EMS PCR Number (TR9.11)"** in the patient registry is rarely (if not at all) utilized by hospitals in that database. The CEMSIS ePCR is the most highly specified unique alpha-numeric identifier that is available to locate an EMS patient, but this element is not accessible or recorded in the trauma patient registry for EMSA to obtain.

However, hospitals use **"Incident Number (TR5.12)"** in the patient registry for multiple patients on different incident dates and different facilities. The NEMSIS data dictionary designates **"Response Incident Number (eResponse.03)"** as the "incident number assigned by the 911 Dispatch System and can be used to associate multiple EMS responses, dispatch information, and other information to the same EMS event or patient." **"Response EMS Response Number (eResponse.04)"** is defined in the NEMSIS data dictionary as the "internal EMS response number which is unique for each EMS Vehicle's (Unit) response to an incident" within an EMS Agency.[iv] These two data elements in CEMSIS were used numerous times for different patients and by different EMS agencies and was deemed unmatchable to hospital patient records. When looking at matched records by incident date and date of birth, these data elements (eResponse.03 and eResponse.04) did not match with any trauma registry elements. Additional investigation into how hospitals use these identifiers is needed.

Furthermore, there were no other unique identifiers in either database that linked to each other. EMS and patient registry logged times were sparse and inconsistent across the two databases, so elements like time delivered/transferred to hospital and admit to ED date/time were not viable for comparison.

Using deterministic research theory, EMSA attempted to match patient records using date of incident and date of birth. The following are the additional criteria and data elements that were used to query the sample population to show further insight into each patient across a continuum of care. These extra elements are not exhaustive but will help demonstrate what type of data analyses are possible for further research studies.

Outline: EMS and Trauma Variables

1. Selected one trauma-level hospital in Riverside County: Riverside Community Hospital (level I trauma center)
2. Selected timeframe: January 1, 2019 through June 30, 2019
3. Selected inclusion criteria for CEMSIS transactional report:
    a. Incident date is not blank

b. Date of Birth (DOB) is not blank
c. Data elements:
    i. Incident date
    ii. Patient date of birth (ePatient.17)
    iii. Patient age (ePatient.15)
    iv. Patient gender (ePatient.13)
    v. Patient race (ePatient.14)
    vi. LEMSA
    vii. Response EMS agency (eResponse.02)
    viii. Situation primary complaint statement (eSituation.04)
    ix. Situation primary provider impression (eSituation.11)
    x. Situation possible injury (eSituation.02)
    xi. Situation initial patient acuity (eSituation.13)
    xii. Situation primary symptom (eSituation.09)
    xiii. Disposition EMS transport method (eDisposition.16)
    xiv. Disposition transport mode from scene (eDisposition.17)
    xv. Disposition reason for choosing destination (eDisposition.20)
    xvi. Cause of injury (eInjury.01)
    xvii. Injury trauma center criteria list (eInjury.03)
    xviii. Incident dispatch notified time (eTimes.02)
    xix. Incident unit notified by dispatch time (eTimes.03)
    xx. Incident unit en route time (eTimes.05)
    xxi. Incident unit arrived on-scene time (eTimes.06)
    xxii. Incident unit left scene time (eTimes.09)
    xxiii. Incident unit patient transfer of care time (eTimes.12)
    xxiv. Response type of turnaround delay (eResponse.12)
    xxv. Response type of scene delay (eResponse.10)
    xxvi. Response time: incident unit arrived on-scene minus incident unit notified by dispatch
    xxvii. Scene time: incident unit left scene (eTimes.09) minus incident unit arrived on-scene(eTimes.06)
    xxviii. Ambulance Patient Offload Time (APOT): incident destination transfer of care (eTimes.12) minus incident patient arrived at destination (eTimes.11)

4. Selected inclusion criteria for the ImageTrend® Patient Registry transactional report
    a. Incident date is not blank
    b. Date of Birth (DOB) is not blank
    c. Data elements:
        i. Incident date
        ii. Patient DOB

iii. Patient gender
iv. Patient race
v. Patient ethnicity
vi. EMS unit notified time
vii. EMS unit arrived on-scene time
viii. EMS unit left scene time
ix. EMS unit at destination time
x. Injury Severity Score (ISS) calculated
xi. ICD-10 injury description
xii. ICD-10 injury detailed description
xiii. Trauma type with ICD-10 COI codes
xiv. Trauma triage criteria (Steps 1 and 2)
xv. Trauma triage criteria (Steps 3 and 4)
xvi. ED/Acute care disposition
xvii. Hospital discharge disposition
xviii. Facility name
xix. Transport to your facility by (method of transport)
xx. Probability of survival
xxi. Ventilator days-total

5. Reports were generated and exported to Excel as ".csv" files, which were then converted into two ".xlsx" files

Phase Two

Matched and queried both tables using SAS® software

1. Imported both files into SAS® Enterprise Guide
2. Linked the tables using an inner join (only incident date and date of birth criteria from both are used in determining matches)
3. 490 out of 763 hospital records were mapped to EMS records (64% match rate)
4. Exported matched data to Excel to clean and analyze

Phase Three

Cleaned and standardized data in Excel

1. Verified records for accuracy and completeness
    a. Manually entered missing or inaccurate ages, zip codes, etc.
    b. Cleaned up and abbreviated variable names and record names for brevity
    c. Deleted duplicate columns that were unnecessary
    d. Converted dates and times to be consistent
    e. Manually verified if there were more viable matching records

f. Checked for date transpositions, incorrect or misspelled words, insertions, omissions, etc.

RESULTS

Out of 763 trauma records, 490 were successfully matched between prehospital (EMS) and hospital (trauma) databases. This yielded a 64% match rate based on the number of trauma records that were queried from Riverside Community Hospital. The following is a breakdown of additional statistics based on the linkage findings.

Figure 1: Flow Chart of Steps 1 and 2

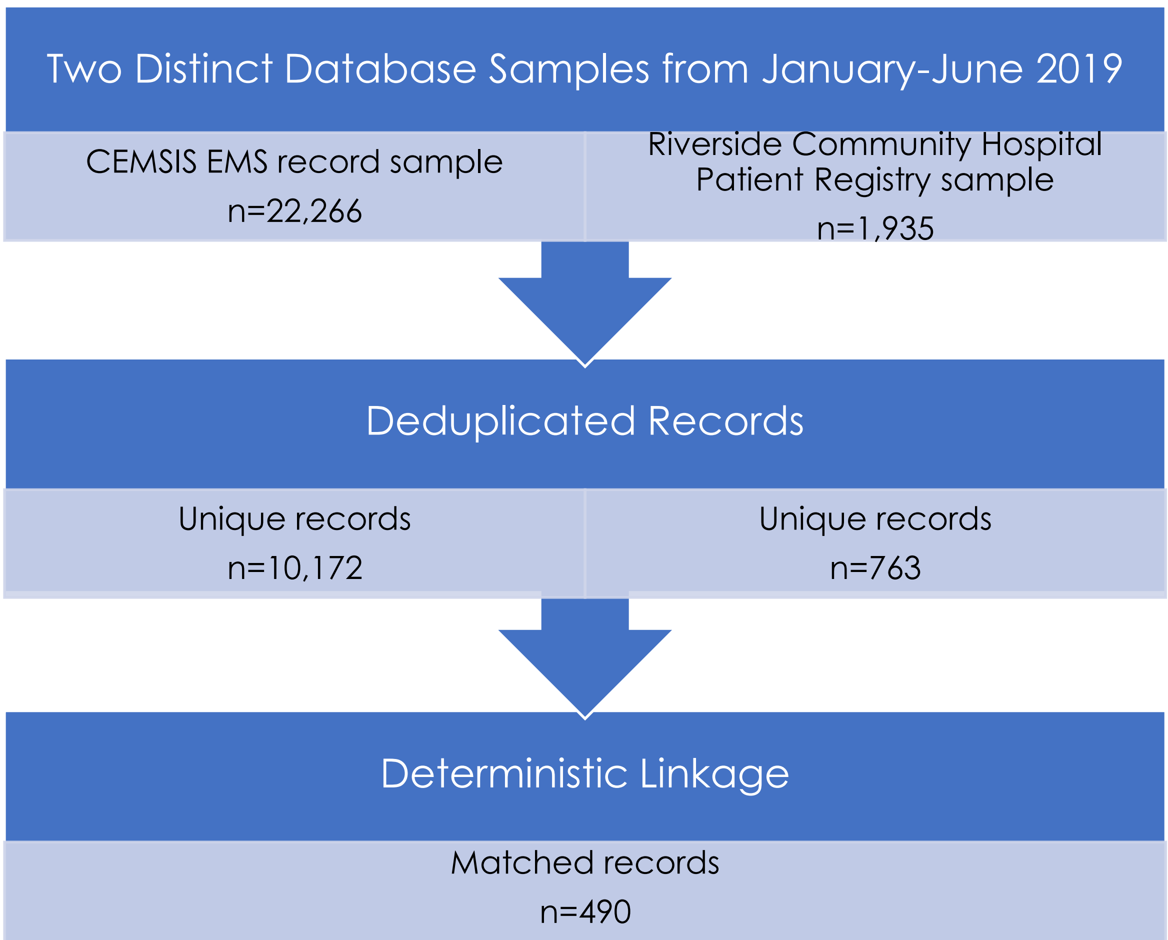
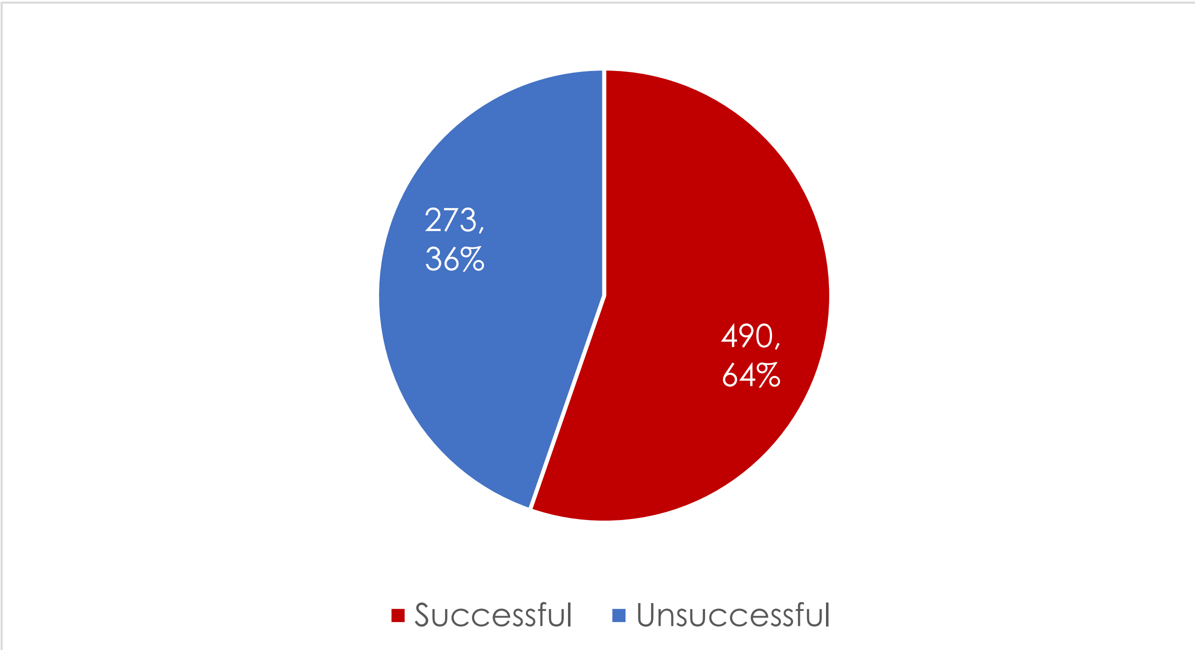


Figure 2: Number and Percentage of Successful and Unsuccessful Matched Records



Potential Descriptive Statistics and Demographic Representations from Data Linkage

- Percent match rates by hospital, hospital system, LEMSA, and EMS provider
- Top primary provider impressions
- Top medications and EMS interventions utilized by type of patient
- Patients by Trauma Type (ICD-10 Cause of Injury Code) compared to how it is recorded in the trauma registry
- Injury severity score and hospital disposition outcomes
- Reasons for choosing facility
- Trauma triage criteria utilization and destination chosen
    - Rates of under-triage, over-triage, and activation of trauma team
- Response, scene, transport, and offload times by EMS provider and LEMSA
- Types of delays and delays based on scene safety, location, or type of emergency (medical or trauma)
- Primary symptoms for a given condition or specified primary impression
- Mean, median, mode of patient race, ethnicity, age and age range, gender
- Geographic information system (GIS) database mapping

CONCLUSION

Linking patient hospital and prehospital records demonstrates the utility, for healthcare stakeholders, EMS agencies, and other interested entities, in identifying gaps in patient care. EMSA was successful in this endeavor, but the match rate can improve through further investigation.

The successful match rate (64%) is comparable to other studies that EMSA identified and had a 6% higher match rate than was done in the previous data-matching study for UC Davis Medical Center. As indicated earlier in this report, multiple other researchers' attempts at data linkage had similar results with an average of 50-70% match rates using only incident dates and the patient's date of birth. They had additional success using more unique identifiers such as full names and ePCR numbers in subsequent matching attempts, which may aid in higher match rates for quality improvement research. It is also worth mentioning that various EMS run times were not utilized due to the inconsistency in which they were recorded in both databases. They could not be verified in the patient registry records because they were either missing (55%) or did not match the EMS record times. Further research into data record completeness is needed to overcome the matching limitations demonstrated in this report.

Future Considerations:

- Conduct a more thorough deterministic/probabilistic study to capture more matches and broaden the study population to other hospitals, LEMSAs, etc. throughout California.
- Match and display data geospatially. One example of a research report is to see the correlation of rural versus urban area time-related and time-sensitive health outcomes such as STEMI and stroke conditions.
- Analyze unmatched EMS and trauma records to find and correct data discrepancies.
- Open discussions with other hospitals, LEMSAs, and government agencies to coordinate data sharing for further research and statistical analysis.
- Explore and present the benefits for these stakeholders to cooperate and create outcomes data they are interested in (i.e., stroke, STEMI, motor vehicle collisions, etc.).
- Conduct comparative analyses based on national trends, EMSA benchmarks and quality improvement (i.e., patient offload times, over or under-triaged trauma patients, advanced hospital notification for suspected stroke, STEMI, or cardiac patients, or effectiveness of certain medical interventions in EMS settings.

REFERENCES

[i] Blanchard, I. E., Williamson, T. S., Ronksley, P., Hagel, B., Niven, D., Dean, S., Shah, M. N., Lang, E. S., & Doig, C. J. (2021). Linkage of Emergency Medical Services and Hospital Data: A Necessary Precursor to Improve Understanding of Outcomes of Prehospital Care. *Prehospital Emergency Care*, 1–10. https://doi.org/10.1080/10903127.2021.1977438

[ii] Chikani, V., Blust, R., Vossbrink, A., Wightman, P., Bissell, S., Graw, J., Martinez, R., & Fisher, B. (2019). Improving the Continuum of Care by Bridging the Gap between Prehospital and Hospital Discharge Data through Stepwise Deterministic Linkage. *Prehospital Emergency Care*, *24*(1), 1–7. https://doi.org/10.1080/10903127.2019.1604925

[iii] Seymour, C. W., Kahn, J. M., Martin-Gill, C., Callaway, C. W., Angus, D. C., & Yealy, D. M. (2014). Creating an Infrastructure for Comparative Effectiveness Research in Emergency Medical Services. *Academic Emergency Medicine*, *21*(5), 599–607. https://doi.org/10.1111/acem.12370

[iv] NEMSIS Data Dictionary Version 3.5.0. (2019). National Highway Traffic Safety Administration (NHTSA). NEMSISDataDictionary.pdf