Executive Summary

In 2016, the National Academies of Sciences, Engineering, and Medicine (NASEM) released a report entitled *A National Trauma Care System: Integrating Military and Civilian Trauma Systems to Achieve Zero Preventable Deaths After Injury.* In response, the American College of Surgeons (ACS) and the National Association of State Emergency Medical Services Officials (NASEMSO), with support from the National Highway Traffic Safety Administration (NHTSA), sought to provide a recommendation for and direction on how best to link record-level trauma data across emergency medical services (EMS) and trauma hospital encounters "spanning the entire continuum of care". We propose a deterministic approach relying on the introduction of a universally unique and anonymous identifier (UUID) assigned to an emergency medical services (EMS) record. The UUID will be recorded in a matched trauma registry record when a hospital abstractor is completing a trauma registry record, as a primary methodology. This statement provides an overview of the benefits and considerations of this approach with reference to the technical aspects, privacy, accessibility, and data quality. Further, the merits of a UUID are outlined with respect to scalability and the longer term strategy of linking all encounters from the initial assessment in the field by EMS through the initial receiving center, trauma hospital, rehabilitation and/or other forms of post acute care.

Introduction

This policy statement has as its principal objective, the goal of outlining the foundation for linking pre-hospital trauma data to hospital data at the local, state, and national level. It is derived from a working group representing experts from the ACS and NASEMSO through a project funded by NHTSA. The impetus for this project was the recommendations outlined in the NASEM Report "*A National Trauma Care System – Integrating Military and Civilian Trauma Systems to Achieve Zero Preventable Deaths After Injury*" as outlined in Table 1. While this policy statement is focused on data linkage across pre-hospital and trauma hospital registries, the recommendations are far more encompassing and include the entire continuum of care to rehabilitation and long term functional outcomes.

Table 1: Relevant recommendations and actions from the National Academies of Sciences, Engineering and Medicine Report: *A National Trauma Care System – Integrating Military and Civilian Trauma Systems to Achieve Zero Preventable Deaths After Injury*"

Recommendation:

- The Secretary of Health and Human Services and the Secretary of Defense, together with their governmental, private, and academic partners, should work jointly to ensure that military and civilian trauma systems collect and share common data spanning the entire continuum of care. Within that integrated data network, measures related to prevention, mortality, disability, mental health, patient experience, and other intermediate and final clinical and cost outcomes should be made readily accessible and useful to all relevant providers and agencies (Recommendation 5)

This recommendation has six specific actions:

- Congress and the White House should hold the DoD and the VA accountable for enabling the linking of patient data stored in their respective systems, providing a full longitudinal view of trauma care delivery and related outcomes for each patient.
- The Office of the National Coordinator for Health Information Technology should work to improve the integration of prehospital and in-hospital trauma care data into electronic health records for all patient populations, including children.
- The American College of Surgeons, the National Highway Traffic Safety Administration, and the National Association of State EMS Officials should work jointly to enable patient-level linkages across the National EMS Information System project's National EMS Database and the National Trauma Data Bank
- Trauma registries should develop mechanisms for incorporating long-term outcomes (e.g., patient-centered functional outcomes, mortality data at 1 year, cost data).
- Efforts should be made to link existing rehabilitation data to trauma registry data.
- HHS, DoD, and their professional society partners should engage the NQF in the development of measures to assess the overall quality of trauma care
- Measures should address the patient experience across the continuum of trauma care, from the point of injury, to emergency and in-patient care, to rehabilitation. These measures should be used in trauma quality improvement programs, including the American College of Surgeons Trauma Quality Improvement Program (TQIP).

Background, Benefits, and Challenges

The implementation of organized systems of trauma care over the last 50 years has been associated with a significant reduction in injury-related mortality. These improvements in outcome are largely attributed to progressive EMS systems, designated trauma hospitals, and field trauma triage criteria to ensure that patients receive the right care in the right place at the right time. Further improvements in outcome will arise because of improvements in access to trauma care and advances in both EMS and in-hospital trauma care. Together, this progress can only come about via a complete understanding of the current state and by creating a "learning health system" associated with trauma care. In such a system, data and experience are systematically integrated with external evidence and ultimately that knowledge is put into practice to advantage patients[1].

While we speak of integrated systems of trauma care, the reality is that the two core infrastructure elements – EMS agencies and trauma hospitals--function as silos in most jurisdictions. The capacity for learning as a system is limited as there is minimal sharing of data at either the patient or system level. Over the last several years, EMS registries and in-hospital trauma registries have standardized their data fields such that all EMS agencies and trauma hospitals capture their data in a standardized manner consistent with their respective data dictionaries as defined by National EMS Information System (NEMSIS) for EMS data and the National Trauma Data Standard (NTDS) for trauma hospital data. This standardization of data acquisition and electronic record submission opens the potential for comparative effectiveness studies within each phase of care. These opportunities have come to fruition for trauma hospitals in the form of the ACS Trauma Quality Improvement Program (TQIP) but have not moved forward for EMS because the field does not have access to readily available patient outcome data on a large scale[2].

As such, the potential advantages of data linkage are significant. For example, in the current environment, individual EMS providers may receive little or no information about the outcomes of their patients and, thus, very little learning is possible. At the system level, lack of data availability across the continuum of care from EMS dispatch to hospital outcome (trauma hospital or otherwise) challenges our ability to understand whether triage practices and the selection of destination hospitals are consistent with guidelines and where opportunities for improvement might exist. Lastly, there is a wide variety of practices across EMS agencies and regions with variation in how resources are deployed, in provider skill mix, and in policies related to and propensity for field interventions. Absent linking these EMS practices to patient outcomes, it is impossible to apply focused empiricism – identifying what works and what doesn't work and refining practices over time with the goal of continuous quality improvement.

There have been limited studies linking patient-level EMS data to trauma hospital data using existing state registries. The most sophisticated are derived from a research consortium requiring multiple IRBs across many EMS agencies and hospitals. With sound methodology,

match rates just exceed 80% if the most complete EMS data are available[3,4]. For the average EMS agency, where variables might be missing, compromising the success of probabilistic linkage, match rates are in the order of only 50-60%[5]. To date, there have not been robust methods to link these data at the national level. Successful linkage efforts will need to establish a clear benefit with a high rate of success, propose a scalable infrastructure which satisfies HIPAA regulations, and mobilize support for large-scale implementation. This policy statement aims to set the framework for achieving those objectives.

Data Linkage Approaches

There are two primary approaches to data linkage between existing datasets: probabilistic and deterministic. In this document, we do not address topics such as machine learning techniques, fuzzy matching, or other variations, as these are variations of probabilistic linkage.

Probabilistic linkage is used to facilitate linkage across datasets when few (or no) unique person identifiers are available. This process relies on using a wide range of person and incident attributes that alone are not unique, but if used together provide varying degrees of probability that two records are a match[6]. Records that achieve a match probability above a certain threshold are considered adequate matches, and the remaining records are categorized into possible matches and non-matches depending upon methodological preferences. This approach is especially useful if data are being linked between datasets with a tangential relationship, and as such, often provide a more direct connection between datasets than might otherwise be expected. This approach has the ancillary benefit of maintaining a margin of privacy between the datasets, yet allowing linkage to occur. However, this approach is limited by the availability of person and incident attributes, incomplete data in either dataset, and the inadequacies of assuming a "true" linkage derived from statistical inference based on the available identifiers. Due to these limitations, we do not advocate for this approach and the rest of this policy statement is instead focused on deterministic data linkage.

Deterministic linkage represents a direct case matching approach that relies on the existence of unique identifiers being present in each of the independent datasets to be linked. Examples of patient identifiers could include name, date of birth, gender, home address, mobile phone number, or for greater accuracy, social security number. It is also possible to utilize identifiers that uniquely identify a record, such as a Universally Unique Identifier (UUID)[7]. Deterministic methods commonly use rule-based or heuristic-based decision algorithms and are highly dependent on the quality and intrinsic discriminability of the elements common to independent datasets. Deterministic linkage offers a more reliable method for definitively linking records between and among independent datasets, and while the major concern with deterministic linkage is related to privacy and data security, that concern is managed by introducing data elements into related datasets, which have as their sole purpose, facilitating data linkage.

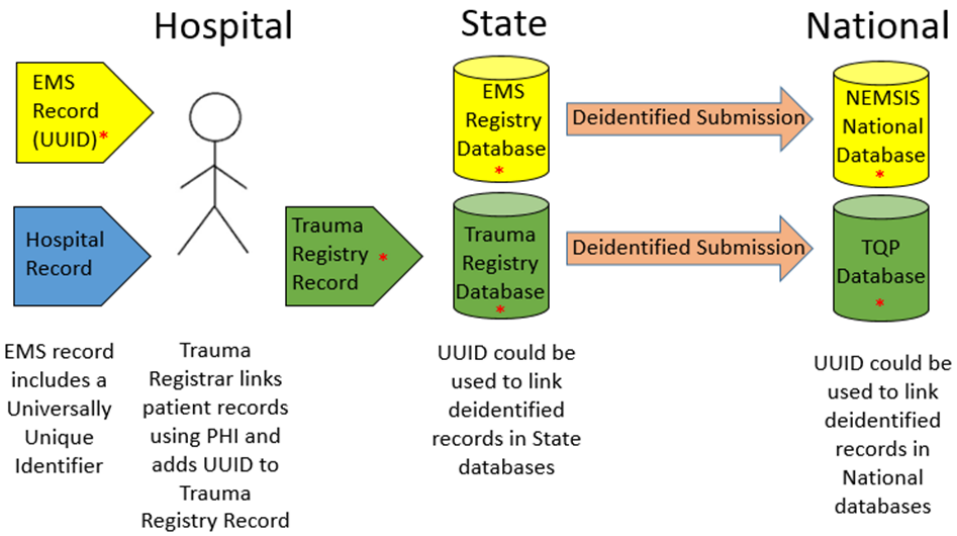
The most effective forms of deterministic linkage are those built on agreed upon unique identifiers. The Social Security Number (SSN) represents a U.S.-based unique identifier, but it has broad relevance outside of health care, and thus, poses a risk to privacy. Further, in the

context of injury, it is rarely available. Instead, we propose that datasets be linked using a procedurally generated Universally Unique Identifier (UUID), or a unique and anonymous record identifier, which facilitates accurate and reliable record linkage without use of identifiable data elements related to a patient or provider. This policy does not detail the technical specifications of a UUID as that information is available elsewhere [8], nevertheless, for the reasons previously stated, *we advocate for deterministic linkage built upon UUIDs as the most effective method for facilitating data linkage between EMS and trauma hospital datasets.*

Deterministic Data Linkage Between EMS and Trauma Hospital Data

Linking record-level data between EMS and trauma hospital datasets can best be achieved by adding the same UUID into each independent dataset and defining business rules for the appropriate transmission, technical and otherwise, of the UUID to national repositories at NEMSIS and the ACS Trauma Quality Programs (TQP). Adding a UUID to an EMS record would allow hospital trauma registrars, abstracting EMS data for trauma registry records, to utilize patient personal health information (PHI) to confirm an accurate match between EMS and hospitals records and transcribe the UUID into the trauma registry record for that patient. The hospital could then transmit the deidentified, linked data to State, Regional, and National EMS and trauma registries (Figure 1). These registries would then include linked EMS and trauma patient data, and there would be no need for transmission of PHI to accomplish national linkage.

Figure 1: Example use case for transmission of an EMS UUID to facilitate anonymous deterministic linkage at State and National levels



Implementation advantages:

1. The initial "linkage" between an EMS record and a trauma registry record is completed by a trained registrar with full access to patient PHI in both records.
2. Local hospital, state, and national registries now all have the same unique and anonymous link between EMS and hospital records for trauma patients, providing a nationally-standardized approach to linkage.
3. The UUID has no intrinsic relationship to a patient, provider or facility.
4. The EMS UUID would be "optional" for submission to ACS TQP database.

We appreciate there may be multiple ways to complete the exchange of the EMS UUID from the EMS record to the trauma hospital record. However, the complexity of the UUID would necessitate an electronic exchange, after a registrar has confirmed the record match, to ensure accuracy. The appropriate exchange strategy will depend on the capacity of existing software and is not specified here, but a generalized use case can be described.

We should note that the NEMSIS project is currently adding a UUID to Version 3.5.0 of the NEMSIS standard (see: https://nemsis.org/v3-5-0-revision/what-is-in-the-revision/uuid/). Also of note, the American College of Surgeons is adding the ability to receive a UUID in the NTDS beginning in 2020. With these national standards in place, a general use case may be described.

Scenario: A 5-year old boy is transported from the scene to the nearest pediatric trauma center. He was an unrestrained passenger in a motor vehicle collision occurring on the corner of State and Main Street. The third day after admission, the hospital trauma registrar begins a record for the boy. Using PHI information, found in the electronic hospital record, she locates the boy's electronic EMS Patient Care Record (PCR). Once she confirms both records relate to the same patient and same event, she "cuts and pastes" (or electronically transfers) the EMS UUID from the NEMSIS compliant EMS PCR to the NTDS compliant trauma registry record. At this point, the same UUID exists in both medical records for the boy, and can be used to deterministically match these independent records at the hospital level, State level and National level if records are exported to the state and national registries (i.e., the NTDB and the National EMS Database).

We believe this approach is the best strategy to link EMS records to trauma hospital records with the goal of creating a learning health system, maintaining privacy, and allowing for continuous quality improvement.

Implications and Important Considerations for Data Linkage

In addition to establishing the infrastructure for UUID-based deterministic data linkage, appropriate implementation of a linking strategy must take into account the following considerations.

Data Privacy

The Health Insurance Portability and Accountability Act (HIPAA) and its implementing regulations (together known as the "Administrative Simplification" provisions) support protections for the privacy of personal health information maintained by covered entities and thus applies to EMS agencies and hospitals[9]. The disclosure of PHI between covered entities is permitted (but not required) without an individual's authorization for health care treatment, payment, and operational purposes, including quality assessment, improvement activities and competency assurance activities. The Privacy Rule assures that health information is protected while allowing electronic flow of health information needed to provide high quality health care and to protect the public.

When considering deterministic data linkage using an UUID, the risk of "loss of privacy" is greatly diminished since a UUID includes no information specific to a patient, provider, or state. In addition, a UUID is meaningless without access to the both independent datasets to be linked. Nevertheless, state variations in data aggregation or data management policies and practices, as originally outlined in the Privacy Rule, may necessitate the use of business associate agreements to exchange the UUID value.

Data Accessibility

This statement recognizes that the accessibility and appropriate use of data are paramount to achieving the specified goals related to quality improvement and comparative effectiveness research. The development of guidelines on the appropriate use of and access to data should be developed prior to the linked data sets becoming available as these data will have great significance to local, State and National stakeholders, including researchers. Commensurate with the increase demand for access and reporting will be a need for enhanced oversight and designation of access rights. These considerations are important since providing access to linked data to local EMS agencies and trauma hospital stakeholders is required to enable local quality improvement activities.

Additional policies should be developed that set proper standards to assure the quality and integrity of all linked data. These policies should define the roles and responsibilities of those with user rights as it relates to access, retrieval, storage, destruction, and backup to ensure proper management and protection of data. It is also imperative that linked data be exempt from Freedom of Information Act (FOIA) requests, and that the data owner (e.g. State) at its sole discretion, can make reports available.

Access and use policies should define and delineate the roles and responsibilities for data usage and establish clear and defined lines of accountability. Policies should be designed around established best practices for effective data management and protection against both internal and external threats, specifically that of privacy and confidentiality. Policies must address all applicable state and federal laws, regulations, and standards.

Data Quality

The utilization of an UUID will limit the need for quality assessments of linkage results. The use of an UUID make data linkage possible with marginal error[10,11]. However, evaluating data quality is still an important step in assuring the resulting linked dataset provides usable information to facilitate data-driven, evidence-based decision making. At a minimum, there are five components of data quality that should be evaluated in each record: Accuracy, Completeness, Consistency, Uniqueness, and Timeliness.

Accuracy refers to the extent to which recorded data reflect the actual underlying information. In other words, accuracy describes the number of errors (incorrectly recorded responses) in a record. While accuracy is not impacted by linkage or a UUID per se, low accuracy limits the value of the linked datasets.

Completeness refers to the extent to which relevant records are present and the fields in each record are populated appropriately. Partial data collection will produce incomplete information and reduce the value of the linked datasets.

Consistency refers to the need to obtain and use data that are clear and well defined enough to yield similar results in different settings. The NEMSIS and ACS TQP standards enable consistent data collection and we do not anticipate any impact of a UUID regarding consistency.

Uniqueness refers to the objective of capturing data once, without unwanted data duplication, and ensuring its application to all appropriate uses. Duplicate data collection in disparate registries leads to unnecessary costs, introduction of random/systematic error and redundant data analysis and reporting. The use of a reliable linkage through a UUID could provide the additional benefit of eliminating the need for redundant data collection in ACS TQP (e.g. pre-hospital vital signs are collected in NEMSIS and then abstracted into the National Trauma Data Standard (NTDS).

Timeliness refers to the need for timely data availability and use. Even slightly dated data reduces the value of the linked datasets. Quality improvement activities based on linked data require these data to be made available promptly after submission. Absent prompt accessibility, it becomes very challenging to improve the quality of care through iterative (i.e., PDSA – Plan-Do-Study-Act) cycles.

Costs and Cost Savings

Previously defined processes designed to probabilistically link de-identified datasets place a fiscal burden on States (or other entities) attempting to link one healthcare dataset to another. Costs were dependent on the chosen linkage software, existing human resources and expertise, additional data storage, etc. In addition, many probabilistic linkage projects include datasets with similar elements (used for linkage), but disparate element definitions and value choices, requiring resource-intensive mapping to ensure best compatibility for linkage with the smallest reduction in the precision of the resulting record matches.

Local and state application of a UUID would reduce linkage costs by pre-defining an anonymous and unique linkage element and relying on trained hospital trauma registrars, with access to all PHI in both EMS and hospital records, to conduct the actual linkage at the time of record abstraction. Existing human resource allocations, including dedicated trauma hospital registrars should not be overtly impacted by this approach, in that the exchange of the EMS UUID to the trauma registry record should be automated. Additional costs might be incurred due to additional data requests and reporting requirements, which together highlight the increase value of the linked datasets.

Conclusion

In summary, we recommend the development and implementation of a Unique Universally Identifier (UUID) in NEMSIS and the NTDS to facilitate deterministic linkage with hospital trauma registries, in accordance with state EMS policies and protections. The principals outlined in this statement allow for an extension of the idea to include the continuum of care from; the initial receiving hospital, transfer to a trauma center, rehabilitation hospitals, and other post-acute care settings. Together, the fulfillment of the recommended actions will allow for continuous quality improvement in trauma care and enable enhanced comparative effectiveness research.

References

1. Haut ER, Mann NC, & Kotwal RS (2016). Military Trauma Care's Learning Health System: The Importance of Data Driven Decision Making. In Berwick D, Downey A, Cornett E (Eds.), A National Trauma Care System: Integrating Military and Civilian Trauma Systems to Achieve Zero Preventable Deaths After Injury. Washington DC: The National Academies Press.
2. https://www.facs.org/~/media/files/quality%20programs/trauma/tqip/tqip_bibliography.ashx
3. Newgard CD, Malveau S, Zive D, Lupton J, Lin A. Building A Longitudinal Cohort From 9-1-1 to 1-Year Using Existing Data Sources, Probabilistic Linkage, and Multiple Imputation: A Validation Study. Acad Emerg Med. 2018 Jul 3. doi: 10.1111/acem.13512. [Epub ahead of print] PubMed PMID: 29969840.
4. Newgard C, Malveau S, Staudenmayer K, Wang NE, Hsia RY, Mann NC, Holmes JF, Kuppermann N, Haukoos JS, Bulger EM, Dai M, Cook LJ; WESTRN investigators. Evaluating the use of existing data sources, probabilistic linkage, and multiple imputation to build population-based injury databases across phases of trauma care. Acad Emerg Med. 2012 Apr;19(4):469-80. doi: 10.1111/j.1553-2712.2012.01324.x. PubMed PMID: 22506952; PubMed Central PMCID: PMC3334286.
5. Lihua Qu, Kevin Putman, Kelly N. Sawyer, Robert Domeier, Jennifer Fowler & William Fales (2018) Challenges of Using Probabilistic Linkage Methodology to Characterize Post-Cardiac Arrest Care in Michigan, Prehospital Emergency Care, 22:2, 208-213

6. Sayers A, Ben-Shlomo Y, Blom AW, Steele F. Probabilistic record linkage. Int J Epidemiol. 2016 Jun;45(3):954-64. doi: 10.1093/ije/dyv322. Epub 2015 Dec 20. PubMed PMID: 26686842; PubMed Central PMCID: PMC5005943.
7. Schaffer, M., Schartner, P., & Rass, S. (2007). Universally Unique Identifiers: How To Ensure Uniqueness While Protecting The Issuer's Privacy. Security and Management.
8. https://tools.ietf.org/html/rfc4122
9. https://www.hhs.gov/sites/default/files/privacysummary.pdf
10. Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, Goldstein H. (2017) Challenges in administrative data linkage for research. Big Data & Society July-December: 1-12
11. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, et al.(2009) The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research. European Journal of Epidemiology 24: 659–667.