A Novel Approach to Data Linkage

EMS and Trauma Registry Records

NEMSIS
BETTER DATA. BETTER CARE.

NASEMSO

AMERICAN COLLEGE OF SURGEONS
Inspiring Quality:
Highest Standards, Better Outcomes
100+*years*

Published 4.5.2021

Table of Contents

Contents

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 the linkage of record-level emergency medical services (EMS) patient care reports with hospital trauma registry records to better elucidate the "continuum of care" for injured patients.

A deterministic linkage approach is proposed, relying on the introduction of a universally unique identifier (UUID) assigned to an EMS record. The UUID would be generated when completing an EMS electronic patient care report (ePCR) and exported to a matched trauma registry record when a hospital abstractor is completing a trauma registry entry. This document provides an overview of this approach, citing the benefits and considerations of the proposal with special reference to technical, privacy, accessibility, and data quality issues related to the approach. Implementation of this method would result in matched EMS/trauma registry records at the local, state, and national levels.

Introduction

This policy statement has as its principal objective, outlining the foundation for linking pre-hospital EMS data to hospital trauma registry data for injured patients at the local, state, and national levels. This proposal is derived from a NHTSA sponsored working group representing subject matter experts from the ACS and NASEMSO (see Appendix). 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*[1] as outlined in Table 1.

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 in the NASEM Report).	
Number	**Recommendation's Actions**
1	Congress and the White House should hold the Department of Defense (DoD) and the Veterans Administration 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.
2	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.
3	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.
4	Trauma registries should develop mechanisms for incorporating long-term outcomes (e.g., patient-centered functional outcomes, mortality data at 1 year, cost data).
5	Efforts should be made to link existing rehabilitation data to trauma registry data.
6	The Department of Health and Human Services, DoD, and their professional society partners should engage the National Quality Framework in the development of measures to assess the overall quality of trauma care.
7	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[2]. 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 likely occur due to continual 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"[1] 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[3].

While we speak of integrated systems of trauma care, the reality is that the patient care information associated with two core health systems – EMS agencies and trauma hospitals, are collected in isolation and function as "data silos" in most jurisdictions. Thus, the capacity for "learning as a system" is limited due to the 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 enables comparative effectiveness studies within each phase of trauma care. These opportunities have come to fruition for trauma hospitals in the form of the ACS Trauma Quality Improvement Program (TQIP) [3] but have not moved forward for EMS because EMS practitioners do not have access to readily available patient outcome data on a large scale.

As such, the potential advantages of data linkage are significant. For example, in the current environment, individual EMS practitioners may receive little or no information about the outcomes of their patients and, thus, very little feedback regarding the care provided is available. This represents a missed learning opportunity. 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 practitioner skill mix, and in protocols related to (and propensity for) field interventions. Without 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 Intuitional Review Boards (IRBs) across many EMS agencies and hospitals. With sound methodology, probabilistic match rates just exceed 80% if the most complete EMS data are available[5,6]. 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%[7]. 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 these objectives.

Data Linkage Approaches

There are two primary approaches to data linkage between existing datasets: probabilistic and deterministic. This document does not address topics such as machine learning techniques, "fuzzy" matching, or other practices, 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 independent records are a match[8]. 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.

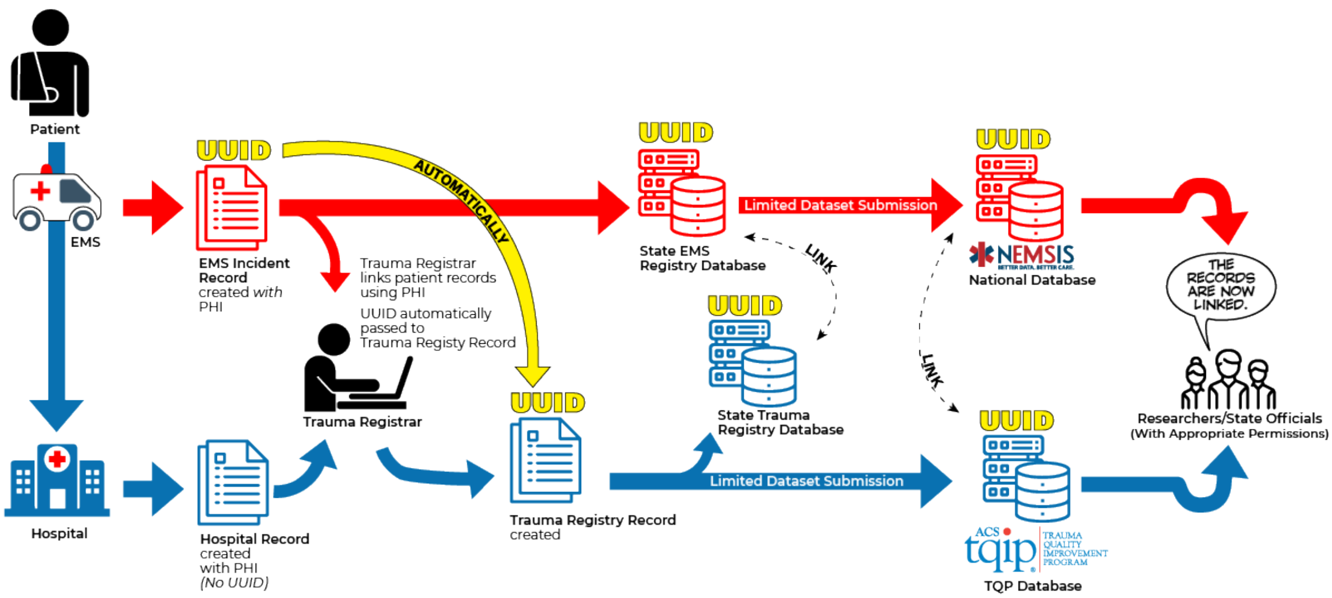
Deterministic linkage represents a direct case matching approach that relies on the existence of unique identifiers being present in each of the independent records to be linked. Examples of patient identifiers could include name, date of birth, home address, mobile phone number, or for greater accuracy, social security number. It is also possible to utilize identifiers that uniquely identify a record, such as a UUID[9]. 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 data linkage, 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 UUID, or a unique and anonymous record identifier, which facilitates accurate and reliable record linkage without the use of identifiable data elements related to a patient or provider. This document does not detail the technical specifications of a UUID as that information is available elsewhere[10], 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 state and national repositories. Attaching 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 hospital records and then automatically export the EMS UUID into the trauma registry record for that patient. The hospital could then transmit records, in compliance with HIPAA, to state and national trauma registries (Figure 1). EMS agencies would also submit de-identified records to the appropriate registries. With the release and uptake of NEMSIS Version 3.5.0, a common UUID will be available to deterministically link EMS and trauma patient data, with no need for transmission of PHI to accomplish state or national linkage.

Figure 1: Example use case for export of an EMS UUID to facilitate deterministic linkage at state and national levels



Implementation advantages: 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.

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. The UUID is a record ID and has no intrinsic relationship to a patient, provider, or facility.

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 has implemented a UUID into Version 3.5.0 using an established standard (see: https://nemsis.org/v3-5-0-revision/what-is-in-the-revision/uuid/). Also of note, the ACS has added the ability to receive a UUID in the NTDS beginning in 2021. With these national standards in place, a general use case may be described.

Scenario: A 5-year old boy is transported from the injury scene to the nearest pediatric trauma center. He was an unrestrained passenger in a motor vehicle collision occurring on Main Street. While on-scene, paramedics completed an electronic Patient Care Report (ePCR), and a UUID associated with the report was automatically generated by the software and stored in the ePCR. On day three of hospital admission, the trauma registrar begins a Trauma Registry record for the boy. Using PHI information found in the two electronic records, she locates the boy's ePCR. Once she confirms both records relate to the same patient and same event, the EMS UUID from the NEMSIS compliant EMS ePCR automatically exports to the NTDS compliant trauma registry record. No PHI is exchanged. At this point, the same UUID exists in both medical records for the boy (i.e., EMS and hospital). The UUID can be used to deterministically match these independent records at the hospital level, state level, and national level, if records are exported to state and national registries (i.e., the National EMS Database and the NTDB/TQIP).

We believe this approach is currently the best strategy to link EMS records to trauma hospital records at the local, state, and national levels 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 consider the following tenets.

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 apply to EMS agencies and hospitals[11]. 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 HIPAA 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 a UUID, the risk of "loss of privacy" is greatly diminished since a UUID includes no information specific to a patient, provider, facility, or state. In addition, a UUID is meaningless without access to both independent records to be linked. Nevertheless, state variations in data aggregation or data management policies and practices, as originally outlined in the HIPAA 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 implemented prior to establishing the linked data sets. Commensurate with an increased demand for access and reporting, there will be a need for oversight and a review 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, privacy, and confidentiality. It is also imperative that linked data be exempt from Freedom of Information Act (FOIA) requests, and that the data owner (e.g., local, state, national levels) at its sole discretion, can make data available. Policies must address all applicable state and federal laws, regulations, and standards.

Data Quality

The utilization of a UUID will increase the quality and accuracy of linkage results between EMS and trauma registry records. That is, the use of a UUID will make data linkage possible with marginal error[12,13]. 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 characterizes the errors (i.e., incorrectly recorded responses) in a record. While accuracy is not impacted by linkage or use of 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 clearly defined to yield similar results in different settings. The NEMSIS and ACS NTDB/TQIP standards enable consistent data collection and will not be impacted by the implementation of a UUID.

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, the introduction of random/systematic error, and redundant data analysis and reporting. The use of a reliable linkage through use of a UUID could provide the additional benefit of eliminating the need for redundant data collection. As an example, the ACS NTDB/TQIP requires the collection of pre-hospital vital signs, which are collected in NEMSIS and then abstracted into the (NTDS). A reliable method of record linkage would reduce the need to "double enter" data in independent registries.

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 are 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 the 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 determine a match among records 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 increased value of the linked datasets.

Conclusion

The utilization of a UUID contained in the EMS record to facilitate deterministic linkage with hospital trauma registries will greatly enhance the efficiency and usefulness of the resulting linkage by limiting the requirement for PHI to utilize probabilistic linkage methods while providing a more complete representation of "continuum of care" information for injured patients at the state and national levels with no exchange of identifying information. The information outlined in this statement allows 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 NASEMO recommended actions (Table 1) will allow for continuous quality improvement in trauma care and enable enhanced comparative effectiveness research.

References

1. Committee on Military Trauma Care's Learning Health System and Its Translation to the Civilian Sector; Board on Health Sciences Policy; Board on the Health of Select Populations; Health and Medicine Division; National Academies of Sciences, Engineering, and Medicine. A National Trauma Care System: Integrating Military and Civilian Trauma Systems to Achieve Zero Preventable Deaths After Injury. Berwick D, Downey A, Cornett E, editors. Washington (DC): National Academies Press (US); 2016 Sep 12. PMID: 27748086.

2. Mann NC, Mullins RJ, MacKenzie EJ, Jurkovich GJ, Mock CN (1999). Systematic review of published evidence regarding trauma system effectiveness. J Trauma, 47(3 Suppl), S25-33.

3. Haut ER, Mann NC, Kotwal RS. Military Trauma Care's Learning Health System: The Importance of Data Driven Decision Making. *National Academies.* 2016.

4. Trauma Quality Improvement Program (TQIP) Bibliography. https://www.facs.org/~/media/files/quality%20programs/trauma/tqip/tqip_bibliography.ashx. Updated November 11, 2018. Accessed May 2018.

5. Newgard C, Malveau S, Zive D, Lupton J, Lin A. A Longitudinal Cohort From 9-1-1 to 1-Year Using Existing Data Sources, Probabilistic Linkage, and Multiple Imputation: A Validation Study. *Academic Emergency Medicine.* 2018;11(25). doi: 10.1111/acem.13512.

6. 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. *Academic Emergency Medicine*. 2018; 19(4), 469-480. doi: 10.1111/j.1553-2712.2012.01324.x.

7. Qu L, Putman K, Sawyer K, Domeier R, Fowler J, Fales W. Challenges of Using Probabilistic Linkage Methodology to Characterize Post-Cardiac Arrest Care in Michigan, Prehospital Emergency Care. *Prehospital Emergency Care.* 2017; 22(2), 208-213. doi: 10.1080/10903127.2017.1362086.

8. Sayers A, Ben-Shlomo Y, Blom AW, Steele F. Probabilistic record linkage. *International Journal of Epidemiology*. 2016; 45(3), 954 – 964. doi: 10.1093/ije/dyv322.

9. Schaffer M, Schartner P, Rass S. Universally Unique Identifiers: How to Ensure Uniqueness While Protecting the Issuer's Privacy. *Security and Management.* 2007.

10. Leach P, Mealling M, R Salz. A Universally Unique Identifier (UUID) URN Namespace. https://tools.ietf.org/html/rfc4122. Published July 2005. Accessed October 2018.

11. United States Department of Health & Humans Services. Summary of the HIPAA Privacy Rule. https://www.hhs.gov/sites/default/files/privacysummary.pdf. Updated May 3, 2018. Accessed October 2018.

12. Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, Goldstein H. Challenges in Administrative Data Linkage for Research. *Big Data & Society*. doi: 10.1177/2053951717745678.

13. Ludvigsson J, Otterbald-Olusson, P, Pettersson B, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *European Journal of Epidemiology*. 2009; 24(11), 659 – 667. Doi: 10.1007/s10654-009-9350-y.

Appendix

Members of the Joint ACS/NAEMSO Data Sharing Expert Panel

Avery B. Nathens, MD, MPH, PhD, FACS
Medical Director, ACS Trauma Quality Programs
Professor of Surgery, University of Toronto
Senior Scientist, Sunnybrook Research Institute
Surgeon in Chief, Sunnybrook Health Sciences Centre
Toronto, ON

Greg Brown
Branch Chief
Arkansas Department of Health, Trauma & Emergency Response
Little Rock, AR

Eileen Bulger, MD, FACS
Chair, ACS Committee on Trauma
Professor of Surgery, University of Washington
Chief of Trauma, Harborview Medical Center
Seattle, WA

Michael Chang, MD, FACS
Chair, ACS COT TQIP Committee
Chief Medical Officer & Associate Vice President for Medical Affairs
University of South Alabama
Mobile, AL

Jimm Dodd, MS, MA
Manager, Stop the Bleed Program
American College of Surgeons
Chicago, IL

Brian J. Eastridge, MD, FACS
Chair, ACS COT Trauma Systems Committee
Professor, Department of Surgery
Division Chief, Trauma and Emergency General Surgery
Jocelyn and Joe Straus Endowed Chair in Trauma Research
University of Texas Health Science Center at San Antonio
San Antonio, TX

Antonio R. Fernandez, PhD, NRP, FAHA
Director of Research, EMS Performance Improvement Center
Assistant Professor, Department of Emergency Medicine
University of North Carolina - Chapel Hill
Chapel Hill, NC

Dia Gainor, MPA
Executive Director
National Association of State EMS Officials (NASEMSO)
Falls Church, VA

Mark Gestring, MD, FACS
Co-Chair, ACS COT Stop the Bleed Steering Committee
Medical Director, Kessler Trauma Center
Chief, Acute Care Surgery Division
Professor of Surgery, Emergency Medicine and Pediatrics
University of Rochester School of Medicine
Rochester, NY

Chris Hoeft, MS
Manager, Data and Report Operations
Trauma Quality Programs
American College of Surgeons
Chicago, IL

Jeff Jarvis, MD, MS, EMT-P, FACEP, FAEMS
Texas State Medical Representative
Governor's EMS & Trauma Advisory Council
Texas Department of State Health Services
Georgetown, TX

Alyssa Johnson, RN, MSN, CEN
Trauma System Program Manager
Montana Department of Public Health & Human Services
Helena, MT

N. Clay Mann, PhD, MS, MBA
Professor, NEMSIS TAC P.I.
University of Utah; School of Medicine; Department of Pediatrics
Salt Lake City, Utah

Holly Michaels, MPH
Manager, Trauma Systems Programs
Trauma Quality Programs
American College of Surgeons

Melanie Neal, MS
Senior Manager, Trauma Quality Programs
American College of Surgeons
Chicago, IL

Kevin Putman, MSA, EMT-P, I/C
Michigan EMS Data Manager
Western Michigan University
Homer Stryker M.D. School of Medicine
Kalamazoo, MI

Daniel Smiley
Chief Deputy Director
Emergency Medical Services Authority
Rancho Cordova, CA