

## A National Assessment of EMS Performance at the Response and Agency Level

Michael Redlener, David G. Buckler, Samuel E. Sondheim, Sai Kaushik Yeturu, George T. Loo, Kevin G. Munjal, Jeffrey Jarvis & Remle P. Crowe

**To cite this article:** Michael Redlener, David G. Buckler, Samuel E. Sondheim, Sai Kaushik Yeturu, George T. Loo, Kevin G. Munjal, Jeffrey Jarvis & Remle P. Crowe (12 Feb 2024): A National Assessment of EMS Performance at the Response and Agency Level, Prehospital Emergency Care, DOI: [10.1080/10903127.2023.2283886](https://doi.org/10.1080/10903127.2023.2283886)

**To link to this article:** <https://doi.org/10.1080/10903127.2023.2283886>



Published online: 12 Feb 2024.



Submit your article to this journal [↗](#)



View related articles [↗](#)





View Crossmark data [↗](#)

---



## A National Assessment of EMS Performance at the Response and Agency Level

Michael Redlener<sup>a</sup>, David G. Buckler<sup>a</sup>, Samuel E. Sondheim<sup>a</sup>, Sai Kaushik Yeturu<sup>a</sup>, George T. Loo<sup>a</sup>, Kevin G. Munjal<sup>a,b</sup>, Jeffrey Jarvis<sup>c</sup> , and Remle P. Crowe<sup>d</sup> 

<sup>a</sup>Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York City, New York; <sup>b</sup>Care2U, New York City, New York; <sup>c</sup>Metropolitan Area EMS Authority, Fort Worth, Texas; <sup>d</sup>Clinical and Operational Research, ESO, Austin, Texas

### ABSTRACT

**Background:** In 2019, the National EMS Quality Alliance (NEMSQA) established a suite of 11 evidence-based EMS quality measures, yet little is known regarding EMS performance on a national level. Our objective was to describe EMS performance at a response and agency level using the National EMS Information System (NEMSIS) dataset.

**Methods:** The 2019 NEMSIS research dataset of all EMS 9-1-1 responses in the United States was utilized to calculate 10 of 11 NEMSQA quality measures. Measure criteria and pseudocode was implemented to calculate the proportion meeting measure criteria and 95% confidence intervals across all encounters and for each anonymized agency. We omitted Pediatrics-03b because the NEMSIS national dataset does not report patient weight. Agency level analysis was subsequently stratified by call volume and urbanicity.

**Results:** Records from 9,679 agencies responding to 26,502,968 9-1-1 events were analyzed. Run-level average performance ranged from 12% for Safety-01 (encounter documented as initial response without the use of lights and siren to 82% for Pediatrics-02 (documented respiratory assessment in pediatric patients with respiratory distress) At the agency level, significant variation in measure performance existed by agency size and by urbanicity. At the individual agency performance analysis, Trauma-04 (trauma patients transported to trauma center) had the lowest agency-level performance with 47% of agencies reporting 0% of eligible runs with documented transport to a trauma center.

**Conclusion:** There is a wide range of performance in key EMS quality measures across the United States that demonstrate a need to identify strategies to improve quality and equity of care in the prehospital environment, system performance and data collection.

### ARTICLE HISTORY

Received 25 May 2023

Revised 5 November 2023

Accepted 12 November 2023

### Introduction

Measuring the quality of care provided by emergency medical services (EMS) and demonstrating the clinical value of this care has been a challenge since the inception of modern EMS. The complexity and variability of service design and delivery, levels of service available and differences in protocols and practice patterns contribute to a widely disparate delivery of care (1–6). Over the last 50 years, advances in biomedical science and out of hospital healthcare delivery science have shaped the protocols and practice of EMS, however, the uniformity of implementation is far from complete (7–9). The development of national evidence-based guidelines has created a framework for this process; however, in contrast to hospital-based or ambulatory practice of medicine, there has been a lack of specific nationally accepted, evidence-based measures to both understand the current practice and support future improvement in EMS care.

Historically, prehospital emergency services metrics have focused on response times and out-of-hospital cardiac arrest survival (10–16). A consensus statement from the 2007 Consortium of U.S. Metropolitan Municipalities' EMS Medical Directors recommended multiple key quality

measures, including associated numbers needed to treat, for six different categories including appropriate treatment and hospital choice for prehospital STEMI cases, benzodiazepine administration for seizures, use of noninvasive ventilation for pulmonary edema and COPD exacerbations, time to application of defibrillator in cardiac arrest, and multiple criteria for trauma cases (10).

In 2015, the National Highway Traffic Safety Administration Office of EMS and the National Association of State EMS Officials partnered to develop standardized quality and performance measures for EMS, known as the EMS Compass initiative. The EMS Compass initiative engaged national stakeholders to develop a broad set of quality measures. In 2018 this work was transitioned to a newly established organization, the National EMS Quality Alliance (NEMSQA), which incorporated a robust measure development process and engaged a broad national EMS stakeholder community and federal partners in EMS to adopt the first set of evidence-based, practical EMS quality measures that could work at large-scale with existing EMS data infrastructure in the United States such as NEMSIS (17).

The first National EMS Quality Measure Set (NEMSQA measures) are 11 evidence-based consensus measures that span

eight key clinical areas across EMS practice (18). Finalized and released in 2019, the NEMSQA measures provide a new opportunity to understand the delivery of quality healthcare in the out of hospital setting. This study aims to describe national performance in these quality measures in a national EMS data set at individual response level and at the EMS agency level. This analysis will enable a novel insight into the practice landscape with regard to the NEMSQA measure set as representation of evidence-based care in EMS at a national level.

## Methods

### Study Design and Setting

In this study, the 2019 NEMSQA quality measures were applied to the 2019 National EMS Information Systems (NEMSIS) research data set. This study was determined not to be human subjects research by the institutional review board at the Icahn School of Medicine at Mount Sinai.

### NEMSQA Measures

Table 1 displays the 11 NEMSQA national performance measures. NEMSQA uses a standardized lifecycle approach in developing and maintaining measures that includes thoroughly researching each measure concept to ensure

grounding in evidence-based guidelines and rationale. Measure specifications were then developed clearly defining components of the numerator, denominator and any exclusions using notation intended to represent syntax used in common statistical packages using NEMSIS data element names (pseudocode). Then, these measures went through a rigorous testing process using the pseudocode definitions to determine validity and reliability. The 11 measures used in this study all passed this rigorous process and were released in their entirety with their technical documentation to the public in 2019. We computed each measure's target population (denominator), quality measure success (numerator) using the Measure Technical Documents and recommended pseudocode (18). We were unable to calculate performance for measure Pediatrics-03b as patient weight is not reported to NEMSIS as part of the national requirement.

### Agency Characteristics

In the NEMSIS research data set, individual agencies were assigned an anonymous "agency code" by the Technical Assistance Center (TAC) at the University of Utah, who oversees use of NEMSIS data for research. This agency code was used to identify individual runs from a particular agency. This masked agency code, however, cannot be associated to a

**Table 1.** National EMS quality alliance national measures and run-level performance.

Measure ID	Description	Total Runs	Overall Run-Level Performance (% [95% CI])
Hypoglycemia-01	Percentage of EMS responses originating from a 911 request for patients with symptomatic hypoglycemia who receive treatment to correct their hypoglycemia.	90,001	69.07 [68.77, 69.38]
Pediatrics-01	Percentage of EMS responses originating from a 911 request for patients with primary or secondary impression of respiratory distress who had a respiratory assessment.	95,230	81.66 [81.42, 81.91]
Pediatrics-02	Percentage of EMS responses originating from a 911 request for patients with a diagnosis of asthma who had an aerosolized beta agonist administered.	20,295	61.1 [60.43, 61.77]
Pediatrics-03b	Percentage of EMS responses originating from a 911 request for patients less than 18 years of age who received a weight-based medication and had a documented weight in kilograms or length-based weight estimate documented during the EMS response.	N/A	N/A
Seizure-02	Percentage of EMS responses originating from a 911 request for patients with status epilepticus who received benzodiazepine during the EMS response.	79,904	26.32 [26.02, 26.63]
Stroke-01	Percentage of EMS responses originating from a 911 request for patients suffering from a suspected stroke who had a stroke assessment performed during the EMS response.	287,119	63.9 [63.72, 64.07]
Trauma-01	Percentage of EMS responses originating from a 911 request for patients with injury who were assessed for pain.	2,766,114	60.77 [60.72, 60.83]
Trauma-03	Percentage of EMS transports originating from a 911 request for patients whose pain score was lowered during the EMS encounter.	1,386,035	15.99 [15.93, 16.05]
Trauma-04	Percentage of EMS responses originating from a 911 request for patients who meet CDC criteria for trauma and are transported to a trauma center.	142,244	21.35 [21.13, 21.56]
Safety-01	Percentage of EMS responses originating from a 911 request in which lights and sirens were not used during response.	26,501,968	11.48 [11.46, 11.49]
Safety-02	Percentage of EMS transports originating from a 911 request during which lights and sirens were not used during patient transport.	15,873,573	53.39 [53.36, 53.41]

geographic area smaller than Census Division nor to identify an individual agency. We defined urban agencies as those who responded to at least 50% of 911-response runs in urban or suburban areas. Run-level urbanicity was assigned in the NEMSIS dataset based on the incident county and the 2013 Urban Influence Codes (19). The NEMSIS research data set contained 4 values for urbanicity: urban, suburban, rural and wilderness. Additionally, we categorized agencies based on overall 911-response in the study year. Categories were created based on prior studies: 1-1,000, 1,001-5,000, 50,01-25,000, 25,001-100,000 and over 100,000 responses (20).

### Statistical Analysis

First, we calculated aggregate run-level proportions of measure criteria met with 95% confidence intervals (CI) for each measure. Next, we calculated overall agency-level performance proportions for each measure and report the median agency-level performance with interquartile range. Visual benchmarking figures for each measure were developed to demonstrate agency-level performance. In this agency-level analysis, individual agency (x-axis) was plotted against performance with both point estimates and 95% CI (y-axis). The overall median performance (all agencies) is represented by the center dotted line with 25% and 75% interquartile ranges added for reference. To better visualize intra-agency variation on each measure, we categorized each agency based on their performance compared to the median agency performance. Agencies with the entire 95% CI below the median line were classified as 'underperforming' (red point). High-performing agencies were those whose 95% CI is entirely above the median line (green point).

To evaluate the association of agency size and the urbanicity of the coverage area with measure performance, we performed additional stratified analyses. As a proxy for agency size, we created groups based on total agency run volume and reported median agency performance and the aggregated percent of total runs each measure accounted for within each size category. We also compared measure performance across urban versus rural service areas and compared median performance proportions between agencies in each stratum. Mood's median tests were employed for each measure to compare performance. As small counts can generate wider variation in estimates, we performed sensitivity analyses at the agency-level including only agencies with at least 10 eligible responses within each measure. All analyses were performed using R Statistical Software version 4.0.2 (21).

### Results

In the 2019 NEMSIS dataset, there were 26,501,968 9-1-1 responses from 9,679 EMS agencies. Included run counts varied considerably across each NEMSQA measure based on the criteria laid forth in each measure definition from a low of 20,295 runs for Pediatrics-02 to a high of 26,501,968 runs for Safety-01. *Measure performance at the run level was observed to have a range of performance from 12% of eligible encounters documented as having a response without the use*

of lights and siren (Safety-01) to 82% of pediatric patients with EMS impressions of respiratory distress having a documented respiratory assessment (Pediatrics-02) (Table 1).

Agency level performance was compared 1) overall, 2) by agency size, and 3) by agency urbanicity. Overall, agency-level measure performance varied widely within and between measures. For example, the median agency performance proportion for Pediatrics-01 was 92% (IQR: 67%-100%), while the median agency performance proportion on Safety-01 was 2% (IQR: 0%-12%) (Figure 1). Results remained similar when limiting analysis to agencies with at least 10 included runs for each measure.

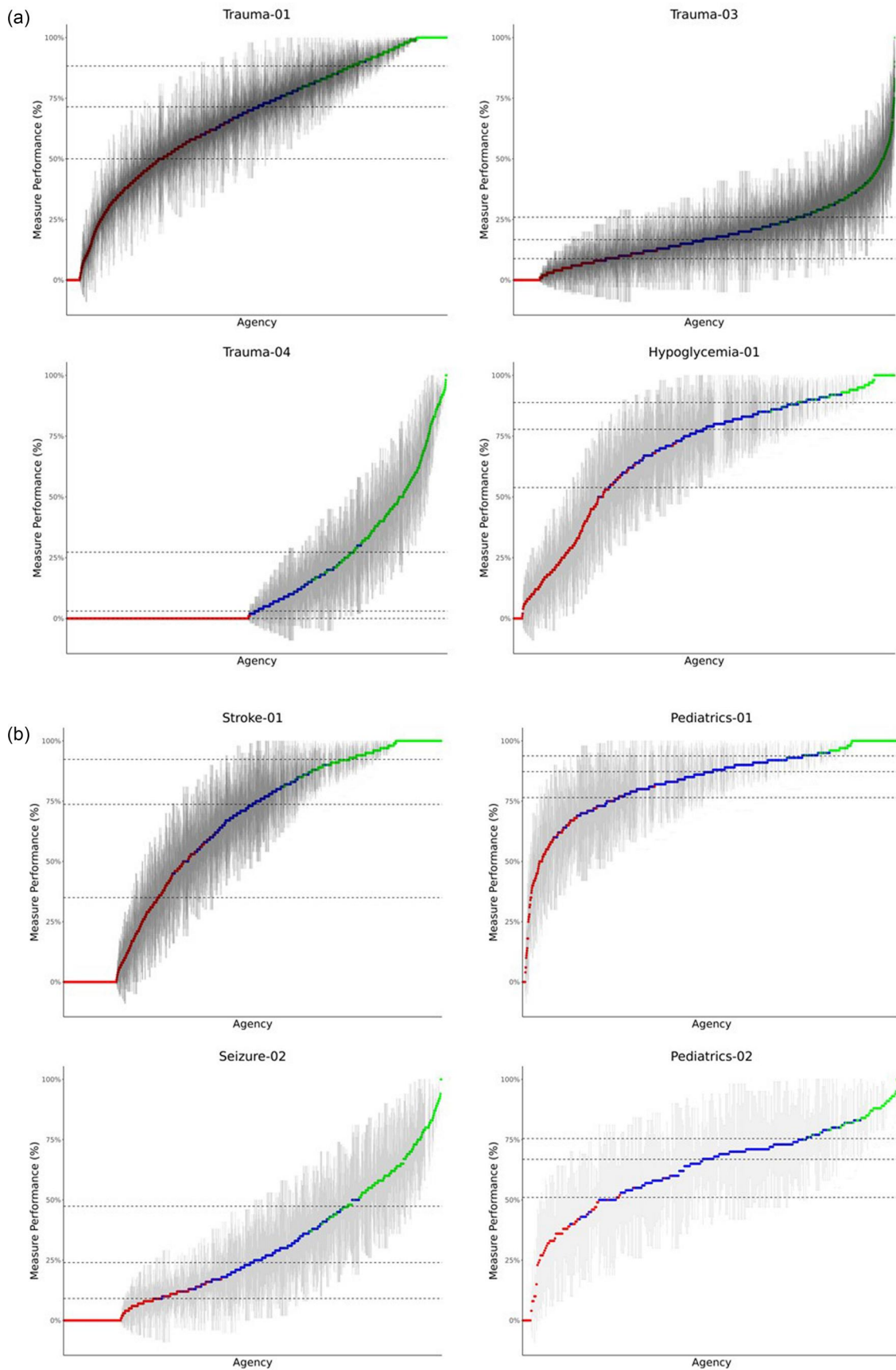
In the analysis by agency size, a total of 22 agencies were included in the largest category with >100,000 runs. Collectively, these agencies accounted for 15% (3,947,678) of all 9-1-1 runs in the year. Meanwhile 6,475 agencies were included in the smallest category 1-1000 total runs. These agencies collectively represented 6% (1,604,162) of the annual 9-1-1 runs. Agency-level measure performance varied by agency size (Table 2). For instance, the median agency-level performance for Hypoglycemia-01 varied from 60% (IQR: 0%-100%) for the smallest agencies to 83% (73%-79%) for agencies with 25,000-100,000 runs.

Substantial variability was also observed in agency-level performance for agencies that primarily responded in rural compared to urban/suburban areas (Table 3). Urban and rural agencies both achieved relative high performance in Pediatrics-01, however, statistically significant differences were observed for six measures. Due to large sample size, most of these differences likely lack practical relevance with the exception of Safety-02 and Trauma-04. Agencies responding in primarily rural areas were less likely to meet or exceed median performance for these measures.

### Discussion

This study is the first description of performance across EMS agencies on a national quality measure set using a national data set. There is clear evidence that EMS care varies widely on important clinical and safety-focused measures. Although benchmarks for acceptable practice have not been developed at the national level, at least half of agencies demonstrated documented performance below 35% for 5 out of the 10 national measures evaluated. Nevertheless, in each measure there were agencies who achieved perfect or near-perfect performance, even after limiting analysis to agencies with at least 10 eligible runs in the measure. While questions remain regarding how data quality and documentation issues impact these findings, this study highlights the variability very clearly for EMS leaders and others concerned about the delivery of care. These findings offer a baseline from which to measure performance in aggregate and track the progress of the national picture.

These multifaceted aggregate and agency-level analyses provide several lenses to consider EMS performance. It is recognized that 100% performance for measures is not always attainable or desirable under the standardized measure definitions and warrant localized attention. While no



**Figure 1.** (1a, 1b, 1c): Agency-Level Performance on NEMSQA National Measures. Each point represents performance for a single agency and the grey whiskers represent 95% confidence intervals (95% CI). Agencies with the entire 95% CI below the median line are ‘underperforming’ (red point). High-performing agencies are those whose 95% CI is entirely above the median line (green point). The central dotted line represents the median performance and the lower and upper dotted lines represent the 25<sup>th</sup> and 75<sup>th</sup> percentile of agency performance. The curve describes the range of performance in each measures across the United States.

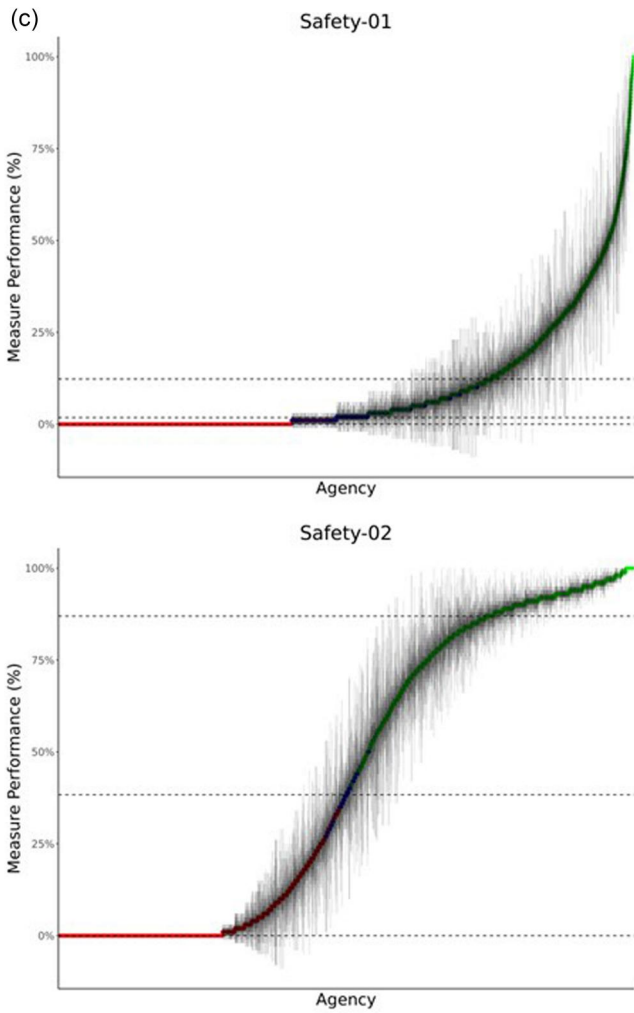


Figure 1. Continued

national targets have been developed for each measure, in some cases, there is research to guide a target. For example, a national analysis of lights and siren use commissioned by the National Highway Traffic Safety Administration provides justification for using benchmark targets of less than 50% lights and siren use for response and less than 5% lights and siren use for transport (22). These targets are further supported by work examining the rate of potentially lifesaving interventions performed by EMS in the context of potential time saved using lights and sirens (23). This study highlighted that a large proportion of responses and transports included use of lights of siren above these targets. The agency-level variation further suggests that lights and siren use may be related to factors at the organization level such as policies and culture. EMS agencies may take this opportunity to reflect on whether there is a specific organizational policy dedicated to lights and siren use or whether a clear protocol should be developed to guide clinicians in this important decision.

For measures where there is more limited information as to a reasonable performance target, the medians and interquartile ranges presented in this study may provide a valuable starting point. For example, agencies may wish

Table 2. Agency-level performance overall and by size.

	Hypoglycemia-01	Pediatrics-01	Pediatrics-02	Seizure-02	Stroke-01	Trauma-01	Trauma-03	Trauma-04	Safety-01	Safety-02
Overall										
Run Count	89,826	95,230	20,295	79,904	2,87,119	27,66,114	13,86,035	1,41,687	2,65,01,968	1,58,73,573
Agency Count	5,302	5,097	2,980	4,279	7,262	7,683	7,167	6,520	9,679	8,430
Performance - Median% [IQR]	75 [29, 100]	92 [67, 100]	58 [0, 100]	15 [0, 50]	75 [25, 100]	73 [50, 91]	15 [5, 25]	0 [0, 20]	2 [0, 12]	33 [0, 87]
Percent of Total Runs % [95%CI]	0.37 [0.36, 0.37]	0.38 [0.38, 0.39]	0.09 [0.09, 0.1]	0.38 [0.38, 0.39]	1.14 [1.13, 1.14]	10.93 [10.92, 10.94]	6.07 [6.06, 6.08]	0.64 [0.64, 0.65]	100 [100, 100]	61.42 [61.4, 61.44]
1-1000 Total Agency Runs										
Run Count	7,392	5,278	1,211	6,300	26,050	187,369	109,909	28,626	1,604,162	981,531
Agency Count	2,407	2,214	859	1,956	4,205	4,647	4,293	3,716	6,475	5,338
Performance - Median% [IQR]	60 [0, 100]	100 [50, 100]	50 [0, 100]	0 [0, 33]	75 [11, 100]	77 [52, 95]	13 [0, 28]	0 [0, 14]	1 [0, 10]	19 [0, 81]
1,001-5,000 Total Agency Runs										
Run Count	21,438	16,556	3,507	18,667	67,975	573,420	307,861	33,066	5,084,023	3,255,424
Agency Count	1,973	1,955	1,327	1,561	2,129	2,106	2,023	1,965	2,234	2,148
Performance - Median% [IQR]	79 [50, 100]	89 [70, 100]	50 [0, 100]	25 [0, 55]	71 [32, 92]	68 [47, 85]	17 [10, 26]	0 [0, 22]	2 [0, 16]	58 [0, 90]
5,001-25,000 Total Agency Runs										
Run Count	29,967	28,390	5,354	25,270	93,151	864,293	434,078	34,458	8,187,884	5,071,530
Agency Count	757	755	630	618	762	763	712	697	796	772
Performance - Median% [IQR]	79 [56, 89]	88 [76, 95]	62 [40, 77]	39 [15, 64]	73 [37, 91]	64 [40, 84]	14 [9, 20]	3 [0, 26]	3 [0, 22]	77 [7, 92]
25,001-100,000 Total Agency Runs										
Run Count	20,224	33,038	5,825	16,757	75,793	823,085	376,695	25,494	7,678,221	4,743,772
Agency Count	144	151	142	124	145	145	118	126	152	150
Performance - Median% [IQR]	83 [73, 89]	87 [77, 93]	69 [56, 78]	55 [35, 73]	80 [49, 93]	61 [14, 94]	14 [9, 18]	0 [0, 38]	6 [0, 22]	87 [17, 93]
> 100,000 Total Agency Runs										
Run Count	10,805	11,968	4,398	12,910	24,150	317,947	157,492	20,043	3,947,678	1,821,316
Agency Count	21	22	22	20	21	22	21	16	22	22
Performance - Median% [IQR]	82 [68, 88]	85 [79, 92]	67 [54, 73]	37 [18, 60]	78 [23, 91]	51 [26, 92]	11 [7, 16]	16 [0, 35]	0 [0, 4]	0 [0, 80]

**Table 3.** Measure performance with urbanicity analysis.

Measure ID	Urban Agencies			Rural Agencies			p-value
	Run Count	Agency count	Performance Median% [IQR]	Run Count	Agency count	Performance Median% [IQR]	
Hypoglycemia-01	71,192	1,387	79 [52, 89]	7,108	291	75 [60, 86]	<0.001
Pediatrics-01	80,541	1,277	87 [77, 94]	3,238	145	87 [73, 94]	0.031
Pediatrics-02	13,573	356	67 [51, 75]	289	12	57 [45, 73]	0.182
Seizure-02	66,536	1,022	25 [9, 50]	3,766	160	23 [8, 36]	0.134
Stroke-01	247,261	2,843	74 [36, 92]	26,702	902	74 [33, 93]	<0.001
Trauma-01	2,485,482	4,137	70 [49, 87]	274,954	2,118	75 [54, 90]	0.482
Trauma-03	1,221,525	3,691	17 [9, 26]	157,808	1,787	17 [8, 27]	<0.001
Trauma-04	108,599	1,761	5 [0, 33]	18,488	690	0 [0, 15]	<0.001
Safety-01	24,074,461	5,783	2 [0, 14]	2,424,661	3,172	2 [0, 10]	0.096
Safety-02	14,237,603	4,807	50 [0, 89]	1,632,798	2,696	21 [0, 82]	0.002

to compare their performance on different measures to the median performance reported here rather than against a specific target goal where such goals do not exist. Additionally, in the case of measures with low performance and lower run volumes highlighted in this study, EMS agencies may consider forming quality improvement collaboratives to pool data and rapidly test change ideas. The results of this study further suggest that as improvement initiatives are developed, deliberate focus should be taken to include rural agencies as disparities in care for those living in rural settings have long been noted (24). While this study did not examine disparities by patient characteristics, future work should also consider local community needs and monitor for equity in efforts to improve performance.

While the performance noted for measures in this study likely reflects a mix of variation in clinical performance as well as documentation practices, individual agencies may use these findings as a launching point to compare their own performance with national data. This data driven self-reflection can empower agencies to identify areas of strength and areas in need of improvement. These findings are intended to ignite quality improvement conversations, rather than serve as an end-point. Agencies may supplement the data elements with data from additional sources and perform overread of the free-text narratives to further quantify clinical practice versus documentation concerns. Additionally, agencies with performance in the top decile warrant further investigation to identify best practices that can be shared and replicated to improve high-quality patient care.

While EMS research is routinely conducted, including using the NEMSIS data set, to evaluate various process measures and outcome measures, researchers often lack consensus on which measures to choose. For example, in two studies each investigating disparities in trauma care, one used on scene time as a process measure while another used transport time (25, 26). Though both of these, and many other measures may be reasonable quality indicators, the use of differing measures makes comparisons across studies difficult even when the NEMSIS dataset is used. The work presented in this study offers benchmarking for those studies to easily identify disparities using validated measures that are highly repeatable and conserved over various time frames and study types. The integration of NEMSQA measures into NEMSIS research therefore creates opportunity for unified

national quality improvement and research efforts, especially as new measures are developed and validated.

Although the National EMS Quality Measures are mostly process measures that rely on evidence-based care, there may be an opportunity to evaluate the performance of these measures as compared to an outcome-based performance assessment such as out-of-hospital cardiac arrest, which has also demonstrated a wide variability across cities and EMS systems. As data integration improves between hospitals and EMS systems here will be opportunities to assess EMS quality measures and outcomes for patients more broadly. This study offers a foundation for future work to utilize EMS quality measure performance to think more specifically about improving prehospital care.

### Limitations

There are many challenges of capturing actual care in an electronic patient care report (ePCR) and therefore in the NEMSIS data set, however, these findings provide a starting point for improvement of quality and documentation of EMS care nationally. The NEMSIS dataset relies on self-entered data by an EMS provider at the time of care of a patient. This clinical data is then captured by ePCR vendor data mapping which translates clinical data into the NEMSIS data set. While NEMSIS does publish recommended mapping lists for key data elements, each vendor may use a different translation process for this data. Additionally, the ePCR free-text narrative may have information about treatment or context of the encounter that is not captured in the discrete data fields of the NEMSIS dataset.

Regional variation of vendor market capture may impact the quality and performance metrics in this data analysis. Additionally, the NEMSIS research dataset does not include the state of origin which brings limitations to the analysis of state-based regulation, protocol and practice. Further, the NEMSIS national dataset does not capture documentation of patient weight, excluding analysis of one of the eleven NEMSQA measures. Lastly, there are challenges related to reliable capture of some data elements, such as hospital designations. These designations are input by the EMS agencies and vendors and may not reflect all active designations for a facility, likely contributing the low observed performance for Trauma-04.

## Conclusion

There is a wide range of performance in key quality measures in EMS across the United States that demonstrate a need to identify strategies to improve both data collection and quality of care in the out-of-hospital environment. These results demonstrate opportunities for improvement in key areas such as lights and sirens use and trauma center transport. Furthermore, the results delineate areas in which national level change may be fruitful, such as lights and siren use, as opposed to areas where agency level or regional intervention would provide a more efficient approach, such as trauma center transport. Most importantly, the results reported here will serve as an important baseline for future studies and reports to identify, with a standardized set of measures, the success of quality improvement measures.

## Disclosure Statement

Michael Redlener is the President of the Board of Directors for the National EMS Quality Alliance. Jeffrey Jarvis is a board member of the National EMS Quality Alliance

## Funding

This work was supported in part through the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai and supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Advancing Translational Sciences. Research reported in this paper was supported by the Office of Research Infrastructure of the National Institutes of Health under award number S10OD026880 and S10OD030463. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## ORCID

Jeffrey Jarvis  <http://orcid.org/0000-0002-1435-5115>  
Remle P. Crowe  <http://orcid.org/0000-0001-9733-9294>

## References

1. Mehta NK, Allam S, Mazimba S, Karim S. Racial, ethnic, and socioeconomic disparities in out-of-hospital cardiac arrest within the United States: now is the time for change. *Heart Rhythm* O2. 2022;3(6):857–63. doi:10.1016/j.hroo.2022.07.009.
2. Bradford JM, Cardenas TC, Edwards A, Norman T, Teixeira PG, Trust MD, DuBose J, Kempema J, Ali S, Brown CV, et al. Racial and ethnic disparity in prehospital pain management for trauma patients. *J Am Coll Surg*. 2023;236(3):461–7. doi:10.1097/XCS.0000000000000486.
3. Garcia RA, Girotra S, Jones PG, McNally B, Spertus JA, Chan PS. Variation in out-of-hospital cardiac arrest survival across emergency medical service agencies. *Circ Cardiovasc Qual Outcomes*. 2022;15(6):e008755. doi:10.1161/CIRCOUTCOMES.121.008755.
4. Eisenberg MS, Horwood BT, Cummins RO, Reynolds-Haertle R, Hearne TR. Cardiac arrest and resuscitation: a tale of 29 cities. *Ann Emerg Med*. 1990;19(2):179–86. doi:10.1016/s0196-0644(05)81805-0.
5. McNally B, Robb R, Mehta M, et al. Out-of-hospital cardiac arrest surveillance -- Cardiac Arrest Registry to Enhance Survival (CARES), United States, October 1, 2005–December 31, 2010. *MMWR Surveill Summ*. 2011;60(8):1–19.
6. Govindarajan P, Lin L, Landman A, McMullan JT, McNally BF, Crouch AJ, Sasson C. Practice variability among the EMS systems participating in Cardiac Arrest Registry to Enhance Survival (CARES). *Resuscitation*. 2012;83(1):76–80. doi:10.1016/j.resuscitation.2011.06.026.
7. Kupas DF, Schenk E, Sholl JM, Kamin R. Characteristics of statewide protocols for emergency medical services in the United States. *Prehosp Emerg Care*. 2015;19(2):292–301. doi:10.3109/10903127.2014.964891.
8. Martin-Gill C, Brown KM, Cash RE, Haupt RM, Potts BT, Richards CT, Patterson PD. 2022 Systematic review of evidence-based guidelines for prehospital care. *Prehosp Emerg Care*. 2023;27(2):131–43. doi:10.1080/10903127.2022.2143603.
9. Taymour RK, Abir M, Chamberlin M, Dunne RB, Lowell M, Wahl K, Scott J. Policy, practice, and research agenda for emergency medical services oversight: a systematic review and environmental scan. *Prehosp Disaster Med*. 2018;33(1):89–97. doi:10.1017/S1049023X17007129.
10. Myers JB, Slovis CM, Eckstein M, Goodloe JM, Isaacs SM, Loflin JR, Mechem CC, Richmond NJ, Pepe PE. Evidence-based performance measures for emergency medical services systems: a model for expanded EMS benchmarking. *Prehosp Emerg Care*. 2008;12(2):141–51. doi:10.1080/10903120801903793.
11. Cummins RO, Chamberlain DA, Abramson NS, Allen M, Baskett PJ, Becker L, Bossaert L, Deloos HH, Dick WF, Eisenberg MS, et al. Recommended guidelines for uniform reporting of data from out-of-hospital cardiac arrest: the Utstein Style. A statement for health professionals from a task force of the American Heart Association, the European Resuscitation Council, the Heart and Stroke Foundation of Canada, and the Australian Resuscitation Council. *Circulation*. 1991;84(2):960–75. Augustdoi:10.1161/01.cir.84.2.960.
12. Mann NC, Dean JM, Mobasher H, Mears G, Ely M. The use of national highway traffic safety administration uniform prehospital data elements in state emergency medical services data collection systems. *Prehosp Emerg Care*. 2004;8(1):29–33. doi:10.1080/312703002788.
13. Dunford J, Domeier RM, Blackwell T, Mears G, Overton J, Rivera-Rivera EJ, Swor R. Performance measurements in emergency medical services. *Prehosp Emerg Care*. 2002;6(1):92–8. doi:10.1080/10903120290938896.
14. Howard I, Cameron P, Wallis L, Castren M, Lindstrom V. Quality indicators for evaluating prehospital emergency care: a scoping review. *Prehosp Disaster Med*. 2018;33(1):43–52. doi:10.1017/S1049023X17007014.
15. O'Connor RE, Slovis CM, Hunt RC, Pirralo RG, Sayre MR. Eliminating errors in emergency medical services: realities and recommendations. *Prehosp Emerg Care*. 2002;6(1):107–13. doi:10.1080/10903120290938913.
16. Moore L. Measuring quality and effectiveness of prehospital EMS. *Prehosp Emerg Care*. 1999;3(4):325–31. doi:10.1080/10903129908958963.
17. What is NEMSIS. NEMSIS. [accessed 2023 May 3]. Available from: <https://nemsis.org/what-is-nemsis/>.
18. National EMS Quality Alliance: National EMS Quality Measure Set. [accessed 2022 Aug 15]. Available from: <https://www.nemsqa.org/nemsqa-measures>.
19. US Department of Agriculture. Economic research service: urban influence codes. 2021.
20. Redlener M, Olivieri P, Loo GT, Munjal K, Hilton MT, Potkin KT, Levy M, Rabrich J, Gunderson MR, Braithwaite SA, et al. National assessment of quality programs in emergency medical services. *Prehosp Emerg Care*. 2018;22(3):370–8. doi:10.1080/10903127.2017.1380094.
21. R Core Team. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2023. <https://www.R-project.org/>



22. Kupas DF. Lights and Siren Use by Emergency Medical Services (EMS): above all do no harm. Washington (DC): National Highway Traffic Safety Administration Office of Emergency Medical Services; 2017.
23. Jarvis JL, Hamilton V, Taigman M, Brown LH. Using red lights and sirens for emergency ambulance response: how often are potentially life-saving interventions performed? *Prehosp Emerg Care.* 2021;25(4):549–55. doi:10.1080/10903127.2020.1797963.
24. Stopyra JP, Crowe RP, Snively AC, Supples MW, Page N, Smith Z, Ashburn NP, Foley K, Miller CD, Mahler SA, et al. Prehospital time disparities for rural patients with suspected STEMI. *Prehosp Emerg Care.* 2023;27(4):488–95. doi:10.1080/10903127.2022.2061660.
25. Escobar N, DiMaggio C, Frangos SG, Winchell RJ, Bukur M, Klein MJ, Krowsoski L, Tandon M, Berry C. Disparity in transport of critically injured patients to trauma centers: analysis of the National emergency medical services information system (NEMSIS). *J Am Coll Surg.* 2022;235(1):78–85. doi:10.1097/XCS.000000000000230.
26. Ordoobadi AJ, Peters GA, Westfal ML, Kelleher CM, Chang DC. Disparity in prehospital scene time for geriatric trauma patients. *Am J Surg.* 2022;223(6):1200–5. doi:10.1016/j.amjsurg.2021.10.031.