



Introducing the Safety Threats and Adverse events in Trauma (STAT) taxonomy: standardized classification system for evaluating safety during trauma resuscitation

Nemanja Baletic^{1,2} · Jamie Riggs³ · Gerald Lebovic^{4,5} · Andrew Petrosoniak^{1,3,4} · Ryan P. Dumas⁶ · Brodie Nolan^{1,3,4} 

Received: 1 April 2022 / Accepted: 15 May 2022 / Published online: 7 June 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany 2022

Abstract

Purpose Adverse events (AEs) during trauma resuscitation are common and heterogeneity in reporting limits comparisons between hospitals and systems. A recent modified Delphi study established a taxonomy of AEs that occur during trauma resuscitation. This tool was further refined to yield the Safety Threats and Adverse events in Trauma (STAT) taxonomy. The objective of this study was to evaluate the inter-rater reliability of the STAT taxonomy using in-situ simulation resuscitations.

Methods Two reviewers utilized the STAT taxonomy to score 12 in-situ simulated trauma resuscitations. AEs were reported for each simulation and timestamped in the case of multiple occurrences of a single AE. Inter-rater reliability was assessed using Gwet's AC1.

Results The agreement on all AEs between reviewers was 90.1% (973/1080). The Gwet's AC1 across AE categories were: EMS handover (median 0.72, IQR [0.54, 0.82]), airway and breathing (median 0.91, IQR [0.60, 1.0]), circulation (median 0.91, IQR [0.72, 1.0]), assessment of injuries (median 0.80, IQR [0.24, 0.91]), management of injuries (median 1.00, IQR [1.00, 1.00]), procedure related (median 1.00, IQR [0.81, 1.00]), patient monitoring and IV access (median 1.00, IQR [1.00, 1.00]), disposition (median 1.00, IQR [1.00, 1.00]), team communication and dynamics (median 0.80, IQR [0.62, 1.00]).

Conclusions The STAT taxonomy yielded 90.1% agreement and demonstrated excellent inter-rater reliability between reviewers in the in-situ simulation scenario. The STAT taxonomy may serve as a standardized evaluation tool of latent safety threats and adverse events in the trauma bay. Future work should focus on applying this tool to live trauma patients.

Keywords Trauma · Adverse events · Resuscitation · Inter-rater reliability

Introduction

Trauma resuscitation is a dynamic and multidisciplinary effort that requires both a technical and non-technical skill set from all team members [1]. Although the components of this sequential process are recognizable, evaluation of

resuscitation itself has proven challenging to standardize [1]. The fluidity of the trauma resuscitation process, variability between providers and centers, the need for simultaneous assessment and time-sensitive interventions are all cited reasons why most adverse events (AEs) occur during the initial phases of trauma resuscitation [1, 2]. AEs are major contributors to trauma related morbidity and mortality. Common AEs include failure to perform therapeutic

Meetings Presented at: Trauma Association of Canada 2022.

✉ Brodie Nolan
brodie.nolan@mail.utoronto.ca

¹ Department of Emergency Medicine, St. Michael's Hospital, Unity Health Toronto, 30 Bond Street, Toronto, ON M5B 1W8, Canada

² Department of Surgery, Division of Plastic and Reconstructive Surgery, Harvard Medical School, Massachusetts General Hospital/Shriner's Hospitals for Children-Boston, Boston, MA, USA

³ Division of Emergency Medicine, Department of Medicine, University of Toronto, Toronto, ON, Canada

⁴ Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

⁵ Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

⁶ Department of Surgery, Division of Burn, Trauma, Acute and Critical Care Surgery, University of Texas Southwestern, UT Southwestern Medical Center, Dallas, TX, USA

or diagnostic measures, disorganization among staff or with equipment, a lack of familiarity with injury patterns, misinterpretation of case complexity, misdiagnosis, and fixation errors [2, 3]. One study found an average of 6.09 AEs occur per fatal trauma and 3.47 AEs can be directly attributed to patient death, while another estimated communication errors occurring in more than 50% of trauma cases [2, 3].

The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) created the “Patient Safety Event Taxonomy”, which provided a structural categorization of AEs. However, the JCAHO tool is not specific to trauma resuscitation and requires in-hospital information to assess degree of harm [4]. The Trauma Resuscitation Using in situ Simulation Team Training (TRUST) study utilized in-situ simulation (ISS) coupled with video review and a modified framework analysis to identify and quantify latent safety threats (LSTs) within trauma resuscitation scenarios [5, 6]. The intent of TRUST was not to provide a generalizable taxonomy of AEs but rather to trial a novel approach to AE identification during ISS hence it is possible that not every AE was identified. The next investigative step called for a refinement of AE themes and sub-themes which would yield a specific AE taxonomy and establish grounds for use in actual clinical scenarios.

Our research team recently completed a RAND/Delphi study resulting in a taxonomy of trauma-specific AEs grouped into the following categories (airway and breathing, circulation, emergency medical service handover, assessment of injuries, management of injuries, procedure related, patient monitoring and access, disposition, and team communications and dynamics) and each AE was assigned a degree of harm classification system (I [no harm] to V [death]) [1]. This novel tool offered a framework for standardized analysis of trauma resuscitations as well as provided a potential foundation for targeted quality improvement (QI) and patient safety initiatives, including video review [1, 2, 7–9]. Utilizing the taxonomy within institutional trauma video review (TVR) programs may allow for reproducible and reliable analyses [1, 10]. This in turn, represents an opportunity to yield more accurate comparisons of trauma resuscitations, more comprehensive understanding of the sequence of events leading to the AE, and more detailed investigations to determine system- and process-level interventions [1, 10]. This taxonomy, however, required additional refinement and testing prior to optimal applicability in simulated and live clinical patient environments.

Our present study further evolved the RAND/Delphi taxonomy by expanding AE descriptions and removing AEs that were lacking in practical application. These efforts yielded the novel 65-metric, 9 category, “Safety Threats and Adverse events in Trauma” (STAT) taxonomy (Table 1).

Objectives

The objective of this study was to evaluate the inter-rater reliability of the new Safety Threats and Adverse events in Trauma (STAT) taxonomy utilizing 12 ISS trauma scenarios.

Methods

Study design

This was a prospective cohort study to identify AEs during high-fidelity simulated trauma resuscitations and evaluate the inter-rater reliability of the STAT taxonomy using 12 video recorded ISS trauma resuscitations from the TRUST study [5, 6]. The institutional research ethics board approved the study (REB ID # 15-046).

Setting, study size and participants

This study was conducted between July 2021 and November 2021. Two expert reviewers (NB and JR) with medical doctorate level training, expertise relevant to acute trauma resuscitation, and expertise in video analysis reviewed the videos from the previously conducted TRUST study described below. Reviewers were granted remote access to the original TRUST study simulation videos and they were asked to apply the STAT taxonomy to all 12 trauma simulation sessions.

The TRUST study was an ISS study performed at a Canadian Level 1 trauma center with approximately 1200 trauma team activations annually (33% with Injury Severity Score > 16) and 75,000 emergency department (ED) visits [5]. The hospital’s two-bed trauma bay served as the location for study [5, 6]. Scenarios were developed and designed from a thematic review of actual trauma cases flagged for morbidity and mortality review [5, 6]. Adverse events, deviations from protocol adherence and unexpected deaths occurring between January 2013 and December 2014 were reviewed and analyzed by two board certified emergency physicians with > 5 years of experience in trauma care and simulation education [5, 6]. They identified recurrent clinical situations that posed recurring threats to patient safety and four simulated scenarios were developed. One scenario was conducted during each of the twelve unannounced simulation sessions (number of times each scenario conducted): surgical airway [3], blunt trauma with massive hemorrhage [4], trauma precipitated by medical event [3] and penetrating injury [2]. All of the clinical equipment, systems and processes used during live trauma resuscitations in the trauma bay were available to the trauma team during the simulations [5, 6].

Table 1 Safety Threats and Adverse events in Trauma (STAT) Taxonomy

| | |
|--|--|
| EMS handover | Management of injuries |
| Failure or delay to activate trauma team | Medication error |
| Inaccurate or incomplete medical history report | Failure to treat hypothermia |
| Team member(s) absent for EMS handover | Failure to apply or incorrect application of pelvic binder in the setting of open book pelvic fracture |
| Patient assessment begins before EMS handover in stable patient | Failure to offer effective analgesia/sedation to patients |
| EMS handover had to be repeated | Failure to reduce fracture/dislocation in setting of pulseless limb |
| Airway and breathing | Failure to provide patient with unique hospital ID or bracelet within 5 min of arrival |
| Failure to identify need for supplemental oxygen | Failure to administer hypertonic saline or mannitol in setting or presumed head injury with lateralizing signs or unilateral pupil deficit |
| Unanticipated loss of airway | Procedure related |
| Unintentional delay in intubation (> 5 min) | Technical errors |
| Unsuccessful intubation attempt | Equipment failure/missing |
| Malpositioned endotracheal tube | Failure to perform an indicated resuscitative procedure |
| Aspiration event | Iatrogenic injury during procedure |
| Ventilator malfunction | Knowledge deficits concerning equipment location |
| Failure to identify need for chest tube | Performing FAST exam interferes with ability to obtain initial IV access |
| Failure to perform surgical airway when indicated | Bodily fluid exposure or needlestick injury to health care team member |
| Administration of paralytics prior to all teams ready | Patient monitoring and access |
| Failure to discuss, anticipate, or treat hemodynamic instability prior to intubation | Inadequate monitoring (ie. loss of telemetry, pulse oximetry for > 3 min) |
| Circulation | Failure of patient monitoring equipment (ie. patient monitor, EtOC2, temperature probe) |
| Failure to obtain peripheral or central venous access within 5 min of first attempt | Oxygen supply runs out |
| Failure to draw bloodwork within 10 min of arrival | Loss of all central/intravenous access |
| Delay of > 10 min to blood product administration (once blood is called for) | Delay in assessment or treatment due to agitated or combative patient |
| Greater than 1L crystalloid bolus given in presumed hemorrhagic shock | Disposition |
| Failure to administer blood products or initiate vasopressors with ongoing shock (SBP < 90) | Delay more than 15 min waiting for CT |
| Failure to activate massive transfusion protocol (if more than 2 units of blood products required) | Delay more than 15 min waiting for OR (if emergent OR) |
| Failure to administer TXA in presumed hemorrhagic shock and injury < 3 h | Transfer to CT scan with hemodynamically unstable patient |
| Failure to give platelets or fresh frozen plasma if > 6 units of blood product given in trauma bay (ie. only RBC given) | Team communications and dynamics |
| Primary resuscitative line is subdiaphragmatic (ie. femoral line, tibial IO) in patient with positive FAST or open book pelvis | Unclear responsibility and roles |
| Assessment of injuries | Patient care activities delayed or not completed due to task overload/ competing priorities |
| Failure to maintain cervical spine precautions (if indicated) | Team member unavailable |
| Failure to get x-rays before departure from trauma bay (if indicated) | Concurrent conversations preventing team leader communication |
| Failure to complete primary survey before departure from trauma bay | Ineffective team leadership/unclear authority of team leader |
| X-ray misinterpreted | Failure to use closed-loop communication |
| FAST misinterpreted | Clinical team members distracted by non-clinical-related tasks |
| Incomplete exposure of patient | Inadequate personal protective equipment |
| Failure to calculate GCS | Trauma team leader leaves position to participate in patient care without delegating interim leader |
| Failure to measure temperature | |
| Failure to assess circulation and function in injured limbs | |

EMS emergency medical services, FAST focused abdominal sonography in trauma, GCS glasgow coma scale, TXA tranexamic acid, SBP systolic blood pressure, RBC red blood cells, IO intraosseous, EtCO2 end tidal carbon dioxide, ID identification, CT computed tomography, OR operating room

Each 15-min simulation was video recorded by wall and ceiling-mounted cameras capturing all areas of clinical care within the trauma bay. A dedicated overhead microphone was installed on the ceiling and handheld voice recorders were placed at workstations to capture conversations away from the patient. These 12 ISS trauma resuscitation video recordings from the TRUST study were then stored on an encrypted drive and made available to investigators participating in the present study.

Sample size calculation

We set an expected Kappa of 0.85 with a margin of error of 0.3 and two-sided alpha of 0.05, which resulted in a sample size of 12 simulation sessions.

Variables

AEs were reported as either occurrence (1) or non-occurrence (0) for each simulation session. As some AEs could occur multiple times in a single simulation session, reviewers recorded the time of occurrence to ensure they were both counting the same AE. The data collection sheet allowed for up to 3 of the same AEs to be recorded in this manner, which resulted in a total of 90 potential AEs per simulated session, from a 65-metric, 9 category taxonomy. No session had more than 3 of the same AE occur.

Data analysis and outcome measures

Individual AE metric agreement

Each reviewer recorded AE occurrence or non-occurrence, and recorded AEs were then compared across the 12 simulated sessions (Table 1). The total number of AEs identified by each reviewer were then tallied and the mean number of AEs per simulated session was calculated.

Total agreement of STAT taxonomy

Total agreement between reviewers was reported as the percentage of agreed AEs over the total number of potential AEs per simulated session. Inter-rater reliability was calculated using Gwet's AC1 analysis. Gwet's AC1's were reported as medians and interquartile ranges (IQR). Additionally, agreement using Kappa was performed when able to be calculated.

Statistical methods

Cohen's Kappa is commonly used to assess inter-rater reliability [11]. However, it is limited in producing a value in the presence of perfect agreement with no variability in scores, thus we elected to use Gwet's AC1 to assess inter-rater reliability [12]. Both Cohen's Kappa and Gwet's AC1 describe an agreement beyond chance (chance-corrected agreement). Gwet's AC1 scores are interpreted as values ≤ 0 as indicating no agreement and 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement [12, 13]. A two-sided 95% CI was additionally computed for each Gwet's AC1 value. *P*-values were not corrected for multiple comparisons as each domain was viewed as independent of one another.

Training and bias mitigation

A training session on the STAT taxonomy with both reviewers (NB and JR) and the principal investigator (BN) was done utilizing another ISS trauma scenario that was not

included in this study. Reviewers were blinded to each other's scores. Refinements to the STAT taxonomy and associated data dictionary were performed throughout the study to clarify terms for reviewers (Appendix A); however, reviewers remained blinded to each other's responses and inter-rater reliability was only performed upon study completion after all 12 scenarios were scored.

Results

Case-based and overall agreement

Using the STAT taxonomy, containing 65 AE metrics across 9 categories, the reviewers evaluated safety and trauma team performance of 12 ISS sessions. The categories with the most identified AEs were emergency medical services (EMS) handover, airway and breathing evaluation, and assessment of injuries. The most common AEs identified included: team members absent for EMS handover (8/12 cases), failure to measure temperature (8/12 cases), trauma team leader leaves position to participate in patient care without delegating interim leader (8/12 cases) and unclear roles and responsibilities (5/12 cases).

The STAT taxonomy yielded an overall total agreement percentage of 90.1% (973/1080) between reviewers (Table 2). Agreement between reviewers according to individual sessions and overall percent agreement of the STAT taxonomy is listed in Table 2. There was a mean of 7.38 AEs identified in each trauma session. Highest agreement between reviewers occurred in session 12 (94.4%) while lowest agreement occurred in session 3 (83%).

Table 2 Reviewer agreement on identified adverse events for each simulation case

| Simulation session | Adverse event agreement between reviewers per case <i>n</i> (%) |
|--------------------|---|
| Session 1 | 76/90 (84.4) |
| Session 2 | 82/90 (91.1) |
| Session 3 | 83/90 (83.0) |
| Session 4 | 75/90 (83.3) |
| Session 5 | 80/90 (88.8) |
| Session 6 | 81/90 (90.0) |
| Session 7 | 78/90 (86.6) |
| Session 8 | 88/90 (97.7) |
| Session 9 | 81/90 (90.0) |
| Session 10 | 81/90 (90.0) |
| Session 11 | 83/90 (92.2) |
| Session 12 | 85/90 (94.4) |
| Total | 973/1080 (90.1) |

Categorical agreement

The 65 AE metrics are grouped into 9 established categories and listed based on the likely sequence of occurrence during a trauma resuscitation. Categorical agreement was measured as a percentage in each of the AE categories with median, IQR, and Q1/Q3 (Table 3). Agreement using Kappa (when able to be calculated) are presented in Appendix B. Individual Gwet's AC1 for each individual AE are presented in Appendix C.

Discussion

This study demonstrates that the STAT taxonomy, when applied to 12 ISS trauma sessions, consistently performed in the "almost perfect" agreement range and yielded a 90.1% overall agreement between reviewers, demonstrating excellent inter-rater reliability. This refined taxonomy has proven to be highly reliable in the in-situ simulation trauma environment, and consequently may serve to assist in targeted patient safety and quality improvement initiatives utilizing live and/or video review modalities for trauma team performance assessment [7, 8, 10, 14]. An accompanying data dictionary (Appendix A) defines each AE metric in greater detail, serving as a reference to facilitate application.

Utilizing an AE taxonomy within a TVR program

TVR programs are increasingly common [7, 15] yet there lacks a standardized review process to define and quantify AEs and safety threats both within and across institutions. The STAT taxonomy, with demonstrable high inter-rater reliability, offers a solution to enhance the reliability and generalizability of AE reviews. Application of the taxonomy during video review analysis may allow for more accurate comparisons of trauma resuscitations at an individual,

institutional and system level. Furthermore, it may improve our understanding of the sequence of events leading to an AE [1]. Repeated occurrences of an AE support the development of a shared mental model among those tasked with QI and system development [1, 16]. For example, the TRUST study investigators were able to utilize the AEs and latent safety threats identified during the recorded ISS sessions to guide training and education to reduce the time to arrival of blood after activation of massive transfusion protocol by 2.5 min (95% CI 0.03–5.08) [16]. Previous studies have demonstrated a reduction in errors after implementing protocols and policies targeting identified safety threats [2, 17]. Such an approach emulates the systems thinking approach that is common across other high-risk industries [1].

Using the STAT taxonomy to drive performance improvement

The categories with highest number of AEs were (1) EMS handover, (2) airway and breathing, and (3) assessment of injuries. Using the example of EMS handover, we can then further define the essential elements of patient history for paramedics to communicate to the trauma team. The high number of AEs linked to EMS handover suggest an important opportunity for improvement and a recent video review analysis of EMS handover supports these findings [18]. By more accurately characterizing the AE, a customized and appropriate mitigation strategy can be trialed, with an opportunity to rerun the scenario (or immediately apply during actual clinical cases) to evaluate the impact and reduction in the AE frequency [18]. In the case of EMS handover, a potential option might be the integration of a standardized handover tool which improves information transmission between the paramedic and receiving team and reduces handover duration [19].

The STAT taxonomy may also be applied broadly to various phases of a trauma resuscitation. Video-recorded debriefing sessions may utilize the STAT taxonomy to communicate strengths and weaknesses of team performance to which educational interventions can then be tailored toward reducing common AEs identified. Clinician educators and simulation programs can use the STAT taxonomy for assessment of learners and trainees. For example, learners could be scored during ISS trauma sessions, allowing them to benchmark themselves alongside other trainees, and provide an objective measure for competency based medical education. The STAT taxonomy therefore offers increasingly specific and detailed feedback on safety and team performance that requires attention or intervention and can drive continuous quality improvement. Lastly, the STAT taxonomy can evaluate safety and performance across centers and regions. By accommodating reliable comparisons in both intra-facility and inter-facility trauma resuscitation, leaders and

Table 3 Summary of Gwet's AC1 agreement for adverse event categories

| Adverse event category | Median (IQR) |
|---------------------------------|-------------------|
| EMS handover | 0.72 (0.54, 0.82) |
| Airway and breathing | 0.91 (0.60, 1.00) |
| Circulation | 0.91 (0.72, 1.00) |
| Assessment of injuries | 0.80 (0.24, 0.91) |
| Management of injuries | 1.00 (1.00, 1.00) |
| Procedure related | 1.00 (0.81, 1.00) |
| Patient monitoring and access | 1.00 (1.00, 1.00) |
| Disposition | 1.00 (1.00, 1.00) |
| Team communication and dynamics | 0.80 (0.62, 1.00) |

EMS emergency medical services

administrators can identify strengths and weaknesses across entire hospital systems.

Limitations of the study

We acknowledge several limitations of this study. This was a single center investigation utilizing two reviewers of similar educational backgrounds and experience levels. Two physician reviewers with variable trauma resuscitation experience may yield different results in the identification of AEs. Additionally, non-physician reviewers may score the taxonomy differently based on alternative perspectives and priorities, a finding we confirmed previously [1]. Some AEs in the STAT taxonomy were not observed and are expected to occur rarely, which may impact inter-rater reliability. Our study utilized 12 in situ simulation sessions to establish the IRR of this taxonomy while future studies are required to demonstrate similar reliability and validity in live trauma resuscitations. However, ISS is well established as a technique to elicit team performance and AEs that parallel the live clinical arena and it is likely that our findings will be similar during live trauma resuscitations [20, 21].

Conclusion

The STAT taxonomy yielded 90.1% agreement and strong inter-rater reliability during ISS trauma sessions. The STAT taxonomy is a promising tool for a standardized evaluation of latent safety threats and adverse events in the trauma bay. It can be applied at several levels to standardize the evaluation of trauma care including local and regional QI initiatives, and multi-center research collaborations. Future studies are needed to assess the validity and reliability of this taxonomy and demonstrate its utility in multi-center ISS scenarios and ultimately actual trauma resuscitation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00068-022-02007-9>.

Acknowledgements We would like to thank Physician Services Incorporated and the St. Michael's Hospital Association Alternate Funds Plan for their financial support for this study. We also thank Melissa McGowan and Richard Wu for their assistance in study coordination.

Author contribution BN, NB, JR, GL, AP, and RD conceived and designed the study. BN obtained research ethics approval. NB, JR, and BN performed data collection. NB, BN and GL analyzed the data. BN and NB drafted the manuscript, and all authors contributed substantially to its revision. BN takes responsibility for the paper as a whole.

Funding BN was supported by Physician Services Incorporated and the St. Michael's Hospital Association Alternate Funds Plan. The funders did not have a role in study design, data analysis, interpretation of the results, writing the manuscript, or submitting the article for publication.

Data availability A de-identified dataset associated with the paper is available upon reasonable request.

Declarations

Conflict of interest There are no conflicts of interest to declare.

References

- Nolan B, Petrosoniak A, Hicks CM, Cripps MW, Dumas RP. Defining adverse events during trauma resuscitation: a modified RAND Delphi study. *Trauma Surg Acute Care Open*. 2021;6(1): e000805.
- Nikouline A, Quirion A, Jung JJ, Nolan B. Errors in adult trauma resuscitation: a systematic review. *Can J Emerg Med Care*. 2021;23(4):537–46. <https://doi.org/10.1007/s43678-021-00118-7>
- Fitzgerald M, Gocentas R, Dziukas L, Cameron P, Mackenzie C, Farrow N. Using video audit to improve trauma resuscitation—time for a new approach. *Can J Surg J Can De Chir*. 2006;49(3):208–11.
- Chang A, Schyve PM, Croteau RJ, O'Leary DS, Loeb JM. The JCAHO patient safety event taxonomy: a standardized terminology and classification schema for near misses and adverse events. *Int J Qual Health C*. 2005;17(2):95–105.
- Fan M, Petrosoniak A, Pinkney S, Hicks C, White K, Almeida APSS, et al. Study protocol for a framework analysis using video review to identify latent safety threats: trauma resuscitation using in situ simulation team training (TRUST). *BMJ Open*. 2016;6(11):e013683.
- Petrosoniak A, Fan M, Hicks CM, White K, McGowan M, Campbell D, et al. Trauma Resuscitation Using in situ Simulation Team Training (TRUST) study: latent safety threat evaluation using framework analysis and video review. *Bmj Qual Saf*. 2021;30(9):739–46. <https://doi.org/10.1136/bmjqs-2020-011363>
- Dumas RP, Vella MA, Hatchimonji JS, Ma L, Maher Z, Holena DN. Trauma video review utilization: a survey of practice in the United States. *Am J Surg*. 2020;219:49–53.
- Dumas RP, Vella MA, Chreiman KC, Smith BP, Subramanian M, Maher Z, et al. Team assessment and decision making is associated with outcomes: a trauma video review analysis. *J Surg Res*. 2020;246:544–9.
- Jung JJ, Jüni P, Lebovic G, Grantcharov T. First-year analysis of the operating room black box study. *Ann Surg*. 2018;271(1):122–7. <https://doi.org/10.1097/SLA.0000000000002863>
- Nolan B, Hicks CM, Petrosoniak A, Jung J, Grantcharov T. Pushing boundaries of video review in trauma: using comprehensive data to improve the safety of trauma care. *Trauma Surg Acute Care Open*. 2020;5(1): e000510.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37–46.
- Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. 2008;61(1):29–48.
- McHugh ML. Interrater reliability: the kappa statistic. *Biochem Medica*. 2012;22(3):276–82.
- Adams-McGavin RC, Jung JJ, van Dalen ASHM, Grantcharov TP, Schijven MP. System factors affecting patient safety in the OR: an analysis of safety threats and resiliency. *Ann Surg*. 2019;274(1):114–9. <https://doi.org/10.1097/SLA.0000000000003616>
- Quirion A, Nikouline A, Jung J, Nolan B. 2021 Contemporary uses of trauma video review: a scoping review. *CJEM*. 2021;23(6):787–96. <https://doi.org/10.1007/s43678-021-00178-9>

16. Gray A, Chartier LB, Pavenski K, McGowan M, Lebovic G, Petrosoniak A. 2020 The clock is ticking: using in situ simulation to improve time to blood administration for bleeding trauma patients. *CJEM*. 2021;23(1):54–62. <https://doi.org/10.1007/s43678-020-00011-9>
17. LaGrone LN, McIntyre L, Riggle A, Robinson BRH, Maier RV, Bulger E, et al. Changes in error patterns in unanticipated trauma deaths during 20 years: in pursuit of zero preventable deaths. *J Trauma Acute Care*. 2020;89(6):1046–53.
18. Nagaraj MB, Lowe JE, Marinica AL, Morshedi BB, Isaacs SM, Miller BL, et al. Using Trauma Video Review to Assess EMS Handoff and Trauma Team Non-Technical Skills. *Prehosp Emerg Care*. 2021;22:1–8.
19. Iedema R, Ball C, Daly B, Young J, Green T, Middleton PM, et al. Design and trial of a new ambulance-to-emergency department handover protocol: 'IMIST-AMBO.' *Bmj Qual Saf*. 2012;21(8):627.
20. Patterson MD, Geis GL, Falcone RA, LeMaster T, Wears RL. In situ simulation: detection of safety threats and teamwork training in a high risk emergency department. *Bmj Qual Saf*. 2013;22(6):468.
21. Wheeler DS, Geis G, Mack EH, LeMaster T, Patterson MD. High-reliability emergency response teams in the hospital: improving quality and safety using in situ simulation training. *Bmj Qual Saf*. 2013;22(6):507.