REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From – To)
13 Jan 2015	Journal Article	November, 2012 – November, 2014

4. TITLE AND SUBTITLE

Assessing trauma care provider judgement in the prediction of need for life-saving interventions

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Amechi N. Anazodo, Sarah B. Murthi, M. Kirsten Frank, Peter F. Hu, Lauren Hartsky, P. Cristina Imle, Christopher T. Stephens, Jay Menaker, Catriona Miller, Theresa Dinardo, Jason Pasley, Colin F. Mackenzie

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
USAF School of Aerospace Medicine
Aeromedical Research Department
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-SA-WP-JA-2014-0034

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2014-3242, 8-Jul-14

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Human judgement on the need for life-saving interventions (LSI) in trauma is poorly studied, especially during initial casualty management. We prospectively examined early clinical judgement and compared clinical experts' predictions of LSI to their later occurrence. Within 10–15 min of direct trauma admission, we surveyed the predictions of pre-hospital care providers (PHP, 92% paramedics), trauma centre nurses (RN), and attending or fellow trauma physicians (MD) on the need for LSI. The actual outcomes including fluid bolus, intubation, transfusion (<1 h and 1–6 h), and emergent surgical interventions were observed. Cohen's kappa statistic (K) and percentage agreement were used to measure agreement among provider responses. Sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) were calculated to compare clinical judgement to actual patient interventions. Among 325 eligible trauma patient admissions, 209 clinical judgement of LSIs were obtained from all three providers. Cohen's kappa statistic for agreement between pairs of provider groups demonstrated no ''disagreement'' (K < 0) between groups, ''fair'' agreement for fluid bolus (K = 0.12–0.19) and blood transfusion 0–6 h (K = 0.22–0.39), and ''moderate'' (K = 0.45–0.49) agreement between PHP and RN regarding intubation and surgical interventions, but no ''excellent'' (K _ 0.81) agreement between any pair of provider groups for any intervention. The percentage agreement across the different clinician groups ranged from 50% to 83%. NPV was 90–99% across providers for all interventions except fluid bolus. Expert clinical judgement provides a benchmark for the prediction of major LSI use in unstable trauma patients. No excellent agreement exists across providers on LSI predictions. It is possible that quality improvement measures and computer modelling-based decision-support could reduce errors of LSI commission and omission found in resuscitation at major trauma centres and enhance decision-making in austere trauma settings by less well-trained providers than those surveyed.

15. SUBJECT TERMS
Clinical judgment, trauma resuscitation, performance measures, human factors, traumacare, decision-assist.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Amechi N. Anazodo
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	SAR	8	19b. TELEPHONE NUMBER (include area code)

ARTICLE IN PRESS

Contents lists available at ScienceDirect

Injury

journal homepage: www.elsevier.com/locate/injury
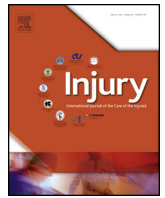
Assessing trauma care provider judgement in the prediction of need for life-saving interventions

Amechi N. Anazodo [a,*], Sarah B. Murthi [c], M. Kirsten Frank [b], Peter F. Hu [a,c],
Lauren Hartsky [b], P. Cristina Imle [a], Christopher T. Stephens [c], Jay Menaker [c],
Catriona Miller [b], Theresa Dinardo [c], Jason Pasley [b,c], Colin F. Mackenzie [a]

[a] Shock Trauma & Anesthesiology Research (STAR) Center, University of Maryland School of Medicine, USA
[b] USAF Center for Sustainment of Trauma and Readiness Skills (C-STARS), Baltimore, MD, USA
[c] R Adams Cowley Shock Trauma Center, University of Maryland Medical Center, USA

ARTICLE INFO

ABSTRACT

Introduction: Human judgement on the need for life-saving interventions (LSI) in trauma is poorly studied, especially during initial casualty management. We prospectively examined early clinical judgement and compared clinical experts' predictions of LSI to their later occurrence.

Patients and methods: Within 10–15 min of direct trauma admission, we surveyed the predictions of pre-hospital care providers (PHP, 92% paramedics), trauma centre nurses (RN), and attending or fellow trauma physicians (MD) on the need for LSI. The actual outcomes including fluid bolus, intubation, transfusion (<1 h and 1–6 h), and emergent surgical interventions were observed. Cohen's kappa statistic (K) and percentage agreement were used to measure agreement among provider responses. Sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) were calculated to compare clinical judgement to actual patient interventions.

Results: Among 325 eligible trauma patient admissions, 209 clinical judgement of LSIs were obtained from all three providers. Cohen's kappa statistic for agreement between pairs of provider groups demonstrated no "disagreement" ($K < 0$) between groups, "fair" agreement for fluid bolus ($K = 0.12$–0.19) and blood transfusion 0–6 h ($K = 0.22$–0.39), and "moderate" ($K = 0.45$–0.49) agreement between PHP and RN regarding intubation and surgical interventions, but no "excellent" ($K \geq 0.81$) agreement between any pair of provider groups for any intervention. The percentage agreement across the different clinician groups ranged from 50% to 83%. NPV was 90–99% across providers for all interventions except fluid bolus.

Conclusions: Expert clinical judgement provides a benchmark for the prediction of major LSI use in unstable trauma patients. No excellent agreement exists across providers on LSI predictions. It is possible that quality improvement measures and computer modelling-based decision-support could reduce errors of LSI commission and omission found in resuscitation at major trauma centres and enhance decision-making in austere trauma settings by less well-trained providers than those surveyed.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Understanding how provider-related decisions impact patient care is critical in efforts to improve care in trauma settings [1]. However, as research and systems development progresses towards the increasing automation of monitoring and decision-assist technology [2–5], particularly to support trauma care in forward-deployed, field, and austere settings, the factoring-in of clinical judgement constants has lagged. Human judgement is complex, influenced by a myriad of social and individual factors, and very difficult to quantify in a manner accessible to math-based systems. In this era of protocol-driven emergency and trauma care, provider judgement is therefore undervalued and understudied. Much of this clinical judgement happens in the background and is particularly uncertain with regard to the anticipation of future

events [6–10]. In preventable trauma deaths, issues with clinical judgement have been found to occur more frequently than those related to skill [11,12]. Additionally, errors of commission and omission leading to trauma mortality still occur even in mature trauma centres [11,13]. Given that most trauma deaths occur within the first 24 h after injury [14], research and development strategies aimed at decreasing trauma mortality must include understanding and improving clinical prognostication as the trigger for timely and effective treatment of potentially fatal injuries.

As part of our ongoing research into human factors in trauma care and developing computer support systems for the next generation of forward-deployable clinical monitoring and decision-assist instrumentation, we asked whether documenting and assessing key prognostic decisions made by three groups of trauma care clinicians—field medical personnel and trauma centre nurses and physicians involved as the patient was being admitted—regarding the proximate need for selected life-saving interventions (LSIs), could provide key insights into the early decision-making process in trauma care. Our goal in this work was assessing the possibility of incorporating such clinically-derived benchmarks into automated clinical instrumentation systems. We therefore undertook a prospective, questionnaire-based study to compare the predictions made by pre-hospital care providers (PHPs), trauma nurses (RNs), and trauma physicians (MDs) in predicting the need for blood transfusion, fluid bolus, intubation, and surgical interventions in critical and unstable trauma patients in the first 24 h of advanced trauma care.

Approval for this prospective, survey-questionnaire-based study was obtained from the University of Maryland, Baltimore and Air Force Institutional Review Boards (IRBs) prior to commencing the study.

Patients and methods

Procedures

After securing the appropriate approvals from the IRBs, we reviewed study procedures and recording instruments with all relevant personnel. These included the dedicated clinical research personnel deployed in the trauma resuscitation unit (TRU) at all times, trauma critical care providers (trauma attending staff, trauma fellows, and specialist registered nursing staff) at our Level I regional adult trauma referral centre and the Maryland State Emergency Medical System (EMS) with which our centre is associated. Informed consent was secured from all TRU attending staff, fellows, and nursing staff. A waiver of the need to document informed consent was approved by all IRBs for the EMS providers. No unique identifying information was collected on the individual PHP other than their years of experience ($\geq$3 years or <3 years) and their status as emergency medical technician (EMT) or paramedic levels of training. Nurses were also asked their years of experience ($\geq$3 or <3 years) on the survey.

All such providers involved in the care of adult ($\geq$18 years old) trauma patients eligible for inclusion in two associated studies [15,16] were eligible to participate in this study. Trauma patients were admitted directly from the scene of injury with a pre-hospital abnormal shock index [SI] $\geq$0.62 (SI = heart rate/minute [HR]/ systolic blood pressure mmHg [SBP]) called in from the field or who were categorised as Priority 1 (critically ill or injured person requiring immediate attention; or unstable patient with life-threatening injury or illness) with or without pre-hospital vital signs per initial field triage. Participating clinicians who had received casualty demographics (age, sex), vital signs, mechanism of injury, and mode and priority of transport called in by radio from the field pre-hospital providers were asked to record, within

10 min of patient arrival, their clinical judgement of the need for LSI likely to be required by that patient at designated intervals within the next 12 h and for blood transfusion up to 24 h. Clinical judgement was recorded via one of three, single-page, pre-validated, survey forms—PHP, RN, or MD (Figs. 1–3, respectively). Forms were collected immediately by research staff and results recorded in the study database.

To avoid the conflict that both the need for the intervention and instituting the intervention were decided upon by the same person, the nurse survey was completed by an experienced nurse not involved in patient care, the field care providers had no input into the in-hospital treatment or decisions about LSI, and the physicians surveyed were consultant (attending) or fellow level in supervisory positions not related directly with details of LSI implementation (such as fluid bolus) that occurred following Advanced Trauma Life Support® guidelines. Providers evaluated the patient based on initial presentation (vital signs, primary survey and EMS history). Providers were then asked to respond on their respective survey forms in yes/no fashion as to whether they thought that any of 12 LSIs listed would be required within the next 12–24 h. Four main LSIs were assessed: fluid bolus, intubation, transfusion, and surgical interventions. Others included cardiopulmonary resuscitation (CPR), medications, use of tourniquets or inflatable splints, etc. Regarding transfusion, three potential timeframes for the requirement for red blood cell transfusion were queried: <1 h, 1–6 h, or 7–24 h. The surgical interventions assessed included emergent surgery to control intra-abdominal haemorrhage, other surgery related to trauma such as pelvic stabilization or other orthopaedic surgery, emergent angiography/embolization, and chest tube insertion. Survey forms were completed by the respective clinicians in separate areas of the TRU and respondents were blinded to the responses of the other subjects.

Fig. 1. Pre-hospital care provider clinical judgement survey form.

ONPOINT CLINICAL JUDGMENT SURVEY

Fig. 2. Nurse clinical judgement survey form.

Clinical research staff then recorded actual LSI outcomes over the first hour of care for each patient and this information was entered into the database. Finally, at 48 h after admission, clinical research staff documented occurrence of relevant interventional and patient outcomes via patient chart review on a fourth survey sheet (Fig. 4) and this information was then recorded in the study database.

Statistical analysis

A convenience sample of eligible trauma admissions was selected. Study data were collected as described above and entered into Access® and Excel® (Microsoft, Redmond, WA) databases for collation and analysis. Statistical assessments were performed using SAS 9.2® (SAS Institute Inc., Cary, NC). Negative and positive predictive values (NPV, PPV), sensitivity (SN), and specificity (SP) were calculated to assess providers' predictions of need for the various LSIs vs. actual occurrence. Cohen's kappa statistic [17] was used to assess the agreement between providers in assessment of the same patient. The percentage agreement (rate of agreement) was also used to compare responses between pairs of provider groups for each intervention. Using the ratings of Landis and Koch, Cohen's kappa values <0.00 indicate poor agreement, 0.01–0.20 indicate slight agreement, 0.21–0.40 are fair, 0.41–0.60 are moderate, 0.61–0.80 are substantial, and 0.81–1.00 indicate excellent agreement [18].

Results

Pilot study data were collected from the providers involved in the care of 50 patients who met admission and injury severity criteria from October to December 2012. After the Air Force IRB

ONPOINT CLINICAL JUDGMENT SURVEY

Fig. 3. Physician clinical judgement survey form.

approval came into effect, a larger study was conducted from February 4, 2013, to May 12, 2013, that included providers involved in the care of 275 additional eligible patients. Of the total 325 study group, all three types of clinicians (PHP, RN, and MD) completed the survey on 209 patients. The remaining 116 eligible patients had less than three provider surveys completed and hence were excluded from analysis.

Demographic and clinical characteristics of these patients are shown in Table 1. Mean Injury Severity Scores (ISS) were 13.0 (SD 13.2) [interquartile range, 4–20]. Injuries included 19% penetrating injuries and 78% blunt injuries (Table 1). Eighty percent of the 209 clinical judgement surveys were completed within 10 min of patient admission to the TRU, the remaining within 15 min. We examined differences in surveys completed 10 vs. 15 min, and in 41 of these surveys (19.6%), it was the physicians who completed the surveys 10–15 min after patient arrival. This was a "real-world clinical study" in which survey completion cannot be mandated. The IRB and the clinicians themselves were insistent that clinical care should take precedence over survey completion. The physicians' greatest involvement was in the initial management, hence limiting their abilities to complete the survey earlier. In addition, the attending physicians and fellows were not necessarily in the TRU admission space itself when the patient arrived but were paged from other clinical duties, thus delaying their timely responses. In a study by Kim et al., among 245 Level I and II trauma centres in the United States, 82% had trauma surgeons available within 15 min of patient admission [19]. Therefore, having survey

2-48 Hour Follow Up for ONPOINT					Study ID:_____
		Ad hoc completed?		_____	
	1.01-3 HOURS	3.01-6 HOURS	6.01-12 HOURS	12.01-24 HOURS	24.01-48 HOURS
CPR (copy code sheet)					
INTUBATION (emergent)					
MONITORS					
Art-line					
CVP/PA					
ICP monitor					
BLOOD (pRBC)					
BLOOD PRODUCTS (FFP or Plateletts)					
Cryoppt					
AUTOTRANSFUSION					
FLUID BOLUS≥1000 ml					
HEMORRHAGE CONTROL PROC.					
Factor VII a					
angio/embolization					
* damage control surgical proc.					
abdominal binder					
tourniquet					
SURGERY RELATED TO TRAUMA					
CHEST TUBE(S)					
PERICADIOCENTESIS					
THORACOTOMY					
HBO					
EXPIRED (DATE & TIME)					
DISPO/DISCHARGE TO: (D&T)					
* Any partial life-saving surgery to "buy time" -- ie. Abdomen packing					
			Data entered: Date_____	Initials:_____	

Fig. 4. 48-h follow-up data collection form–completed by clinical research staff from patient chart reviews.

responses completed by 100% of physician leaders within 15 min and 80% within 10 min would exceed an expected timely survey completion rate at most U.S. trauma centres.

Of the PHP respondents completing the clinical judgement surveys, 92% were paramedics and the remainder were EMTs. More than 90% of the RN respondents had at least 3 years of experience. Among MD survey respondents, (59% surgeons, 31% anaesthesiologists, 6% emergency medicine or critical care), 30% were fellows and the rest were attendings. The surgeon attendings (consultants) had 5–20 years of experience. The fellows had 4–8 years of experience including at least 3 months of intensive casualty reception and resuscitation experience. Three years was the duration suggested by pre-hospital provider EMTs for becoming designated as "experienced." We found no differences between levels of experience of MD attendings vs. fellows, RN ≥ or <3years experience, or paramedics vs. EMTs.

Twenty-two surgical interventions occurred including emergent surgery to control intra-abdominal haemorrhage, pelvic stabilization, orthopaedic surgery, emergent angiography/embolization, and chest tube insertion. Pericardiocentesis was not performed in the 209 casualties included in the survey. One patient had CPR after admission and 11 patients had inflatable splints or pelvic stabilization to control haemorrhage. Except for fluid bolus, NPV was 90% or greater across all providers and categories (Table 2). In contrast, the PPV was greatest for intubation (60–70%) and lowest for surgical intervention (<30%) across providers.

The kappa analysis to determine strength of agreement between providers (Table 3) showed no disagreement in any category (all values are positive). However, agreement on fluid bolus between any two groups of providers was mostly "slight" (0.12–0.19). Agreement on the likely need for transfusion was only "fair" (0.22–0.39) [Make reference to Table 3 for confidence interval ranges]. Nurses and pre-hospital providers achieved "moderate" agreement on the possible need for intubation (0.49, CI: 0.35–0.64) and for surgical intervention (0.45, CI: 0.31–0.59). "Excellent" agreement between the three groups of evaluators was not observed in any category. Overall, the percentage agreement across the different clinician groups ranged from 50% to 83%, with the highest rates across provider pairs observed for intubation (80–83%). On agreement of provider judgement with actual outcome (Table 4), the highest agreement

Table 1
Population characteristics of the 209 patients for whom completed surveys were available.

Characteristic	Value	25th, 75th IQR
Age, year (SD)	41.2 (19.4)	26–52
Male, n (%)	157 (75%)	–
Type of injury, n (%)		
Penetrating	39 (19%)	–
Blunt	159 (76%)	–
Other	11 (5%)	–
Mean ISS (SD)	13.0 (13.2)	4–20
Shock index (SD)	0.74 (0.23)	0.63–0.82

ISS: injury severity score; IQR: interquartile range; SI: shock index: heart rate/systolic blood pressure.

Table 2
Comparison of LSI predictions by trauma clinical care providers.

Intervention	PHP				RN				MD			
	PPV	NPV	SN	SP	PPV	NPV	SN	SP	PPV	NPV	SN	SP
Blood ($n=33$)	0.48	0.90	0.48	0.90	0.34	0.98	0.91	0.68	0.50	0.95	0.53	0.96
Fluid bolus ($n=88$)	0.54	0.61	0.37	0.76	0.49	0.70	0.75	0.43	0.56	0.63	0.40	0.77
Intubation ($n=33$)	0.61	0.94	0.69	0.92	0.60	0.99	0.97	0.87	0.70	0.97	0.87	0.93
All surgical interventions ($n=22$)	0.29	0.95	0.64	0.81	0.27	0.95	0.64	0.80	0.27	0.94	0.54	0.82

LSI: life-saving intervention; PHP: pre-hospital provider; RN: nurse; MD: physician; PPV: positive predictive value; NPV: negative predictive value; SN: sensitivity; SP: specificity.

was observed for nurses ($K = 0.61$, CI: 0.48–0.74), and MD and PHP varied from $K = 0.15$ to 0.48 [Confidence interval ranges shown in Table 4].

Discussion

In this study, we attempted to assess the general agreement between the three main groups of initial trauma care providers (PHPs, RNs, and MDs) in predicting the need for selected LSIs in severely injured trauma patients. Our intermediate goal in this work was increasing understanding of human factors associated with early trauma care decision-making, and our long-term goal was to establish at least suggestive clinical benchmarks against which to assess automated clinical decision-assist systems. The four general categories of interventions assessed (fluid bolus, intubation, transfusion, and surgical interventions) represent different aspects of care, require activation of different team members and elements of the trauma system, and are sampled across a spectrum of severely injured patients. However, cross-overs also existed across these interventions (for example, all of the patients who went to surgery were intubated).

The much lower ability to predict the use of fluid boluses is perhaps not surprising, given that the indications for crystalloid fluid support are less protocolized—that is, its use is more individual provider-based. In addition, at least among trauma physicians, hypotensive resuscitation and decreasing the use of crystalloids have emerged over the last decade as preferred approaches to minimizing blood loss in trauma patients before surgical control of bleeding is achieved, but this is still a focus of active debate [20-22].

With the exception of fluid bolus, for the other intervention categories, all provider groups registered impressive NPVs (90–99%), particularly the nurses and physicians (94–99%), suggesting that knowing who will *not* need an intervention is a critical element of good clinical judgement. Conversely, the PPV was relatively less across the board. Although the overall agreement was similar between PHPs, RNs, and MDs, the Cohen's kappa results suggest that agreement among groups of clinicians is far less dependable.

The overall patterns of agreement among provider groups in our study are not surprising, since all three groups represent unique and consistently highly trained and experienced providers in a mature civilian trauma care system. However, at each level of care, PHP, RN, and MD, subtle interplays of training, experience, and available information must impact prognostication. We did not see "excellent" agreement between the three groups of evaluators in any category of LSI prediction using the kappa analysis. The kappa values and rates of agreement between providers' prediction of the occurrence of these interventions as well as their respective calculated confidence intervals show that when compared against the gold standard (i.e. actual interventions), the agreement among provider groups in prediction of the actual intervention is still far from perfect. These results further support available evidence in the literature that errors of omission and commission can and do occur [11,13].

Since it is well known that even in trauma centres with considerable expertise, decision-support reduces errors of omission and commission during the first 30 min of reception and resuscitation [2]; decision-support tools could perhaps be useful to coordinate the team activities. Although pre-hospital care providers had the most time with each patient through the processes of extraction and transport, they were within the most constrained data-gathering circumstances before being asked to make their prognoses. The attending physicians and fellows, who were not necessarily in the admission space itself when the patient arrived, had the least time with patients before being asked to make their prognoses. Nurses, in effect, had the most advantageous data-gathering and processing circumstances, having seen the patient from the moment of patient arrival and were present when an updated report was called in from the field.

Taken together, our data suggest that expert human judgement is fairly accurate in predicting the need for LSIs among and across highly trained and experienced practitioners. Obviously, training and experience must remain a core focus for patient care improvement among all three groups. However, our results also support the notion that improving the quality and forward-deployability of data-gathering and decision-assist instrumentation has a role.

How human factors research results can best be made accessible to math-based systems developers remains an open question. The confrontation between the classically expert-judgement-based "art of medicine" point of view and that of

Table 3
Agreement between provider groups without reference to outcomes (Cohen's kappa and Percentage Agreement).

Intervention	PHP vs. RN				PHP vs. MD				RN vs. MD			
	Percentage agreement		Kappa		Percentage agreement		Kappa		Percentage agreement		Kappa	
	%	% range CI	K	95% CI	%	% range CI	K	95% CI	%	% range CI	K	95% CI
Blood ($n=33$)	66	59–72	0.22	(0.12–0.34)	75	68–80	0.24	(0.10–0.39)	72	65–78	0.39	(0.26–0.51)
Fluid bolus ($n=88$)	50	43–57	0.12	(0.02–0.22)	67	60–73	0.19	(0.05–0.33)	52	45–59	0.14	(0.04–0.24)
Intubation ($n=33$)	83	77–88	0.49	(0.35–0.64)	83	78–88	0.28	(0.12–0.45)	80	75–86	0.32	(0.18–0.47)
Surgical intervention ($n=22$)	80	74–85	0.45	(0.31–0.59)	78	71–83	0.31	(0.16–0.46)	82	77–88	0.47	(0.32–0.61)

PHP: pre-hospital provider; RN: nurse; MD: physician. Blood: transfusion of packed red blood cells within 6 h of trauma centre admission. Fluid bolus and intubation: within the first hour of care. Surgical interventions: within the first 3 h of care. K: kappa statistic; CI: confidence interval.

Table 4
Agreement of provider judgement vs. actual outcome.

	PHP		RN		MD	
	Kappa	CI	Kappa	CI	Kappa	CI
Blood (n = 33)	0.39	(0.22–0.56)	0.35	(0.24–0.46)	0.33	(0.18–0.48)
Fluid bolus (n = 88)	0.15	(0.02–0.28)	0.17	(0.05–0.29)	0.17	(0.04–0.31)
Intubation (n = 33)	0.48	(0.32–0.63)	0.61	(0.48–0.74)	0.33	(0.15–0.51)
Surgical interventions (n = 22)	0.29	(0.14–0.44)	0.28	(0.13–0.43)	0.25	(0.08–0.42)

PHP: pre-hospital provider; RN: nurse; MD: physician. Blood: transfusion of packed red blood cells within 6 h of trauma centre admission. Fluid bolus and intubation: within the first hour of care. Surgical interventions: within the first 3 h of care. CI: confidence interval.

"evidence-based medicine" is neither the focus nor the concern of the present study; however, the need to merge the two effectively in computer-based support systems is of concern across many health care disciplines, most particularly in critical care, to allow for better triage and distribution of resources [23-26]. Within this general debate, the clearest role for math-based systems development—computer modelling and instrumentation based on this work—is in improved data-gathering and decision-assist tools [27,28]. Our findings support the notion that a specific high-impact area for such research and decision-support development is in pre-hospital transport care, trauma centre patient reception, and resuscitation [2].

The chief limitations in this study are those selection biases classically inherent in any kind of voluntary, convenience-sampling-based, human-subjects research. In this study, these biases operated at several levels. Patient eligibility was objectively based on having to do with the degree of trauma. However, research staff then had to determine whether imposition of even the minimally intrusive non-invasive device would interfere unduly with patient care and then survey results had to be collected from at least one of each kind of provider for each patient. There was some variability in completion of the surveys by physicians in particular that may have allowed them to have knowledge of some of the LSIs; however, this made them no better at predicting them than the other providers. These various levels of bias are likely to have excluded the most severely injured patients. One would anticipate that this would include those with severe, obvious, vascular injury, particularly penetrating injury; however, the proportion of penetrating injury patients was 19%, consistent with the historically expected proportion of 15–30% in our population [29]. Our data do not provide clear clues as to how these biases may have affected outcomes, but the most likely effect is to have damped down rather than inflated agreement across and among providers. An unintended benefit of this effect may have been an added focus on that third of the triage triad who are least clearly visible, that is, neither the worst nor the least injured.

Conclusions

Clinical judgement of expert pre-hospital and trauma centre providers, separately and together, provides a benchmark for the prediction of the need for selected life-saving interventions for patients after severe traumatic injury. The data show that there is no excellent agreement among experts in LSI prediction. Tools comprising computer models and other automated predictors could provide much needed objective data to improve both PPV and agreement across providers. The prediction errors and provider agreement observed in this study may indicate that quality improvement measures and decision-assist tools could reduce errors of LSI commission and omission found in trauma patient resuscitation at major trauma centres and could enhance prompt decision-making capabilities [2] in austere and forward-deployable settings. Further study on the impact of human judgement on outcome is warranted, as well as investigation into tools to augment and improve prediction of life saving interventions.

References

[1] Hoyt DB, Hollingsworth-Fridlund P, Winchell RJ, Simons RK, Holbrook T, Fortlage D. Analysis of recurrent process errors leading to provider-related complications on an organized trauma service: directions for care improvement. J Trauma 1994;36:377–84.

[2] Fitzgerald M, Cameron P, Mackenzie C, Farrow N, Scicluna P, Gocentas R, et al. Trauma resuscitation errors and computer-assisted decision support. Arch Surg 2011;146:218–25.

[3] Clarke JR, Hayward CZ, Santora TA, Wagner DK, Webber BL. Computer-generated trauma management plans: comparison with actual care. World J Surg 2002;26:536–8.

[4] Goud R, de Keizer NF, ter Riet G, Wyatt JC, Hasman A, Hellemans IM, et al. Effect of guideline based computerised decision support on decision making of multidisciplinary teams: cluster randomised trial in cardiac rehabilitation. BMJ 2009;338:b1440.

[5] Moore FA, McKinley BA, Moore EE. The next generation in shock resuscitation. Lancet 2004;363:1988–96.

[6] Christakis NA, Iwashyna TJ. Attitude and self-reported practice regarding prognostication in a national sample of internists. Arch Intern Med 1998;158:2389–95.

[7] Mulholland SA, Gabbe BJ, Cameron P, Victorian State Trauma Outcomes Registry and Monitoring Group (VSTORM). Is paramedic judgement useful in prehospital trauma triage? Injury 2005;36:1298–305.

[8] Perkins HS, Jonsen AR, Epstein WV. Providers as predictors: using outcome predictions in intensive care. Crit Care Med 1986;14:105–10.

[9] Ebell MH, Bergus GR, Warbasse L, Bloomer R. The inability of physicians to predict the outcome of in-hospital resuscitation. J Gen Intern Med 1996;11:16–22.

[10] Kirves H, Handolin L, Niemelä M, Pitkäniemi J, Randell T. Paramedics' and pre-hospital physicians' assessments of anatomic injury in trauma patients: a cohort study. Scand J Trauma Resusc Emerg Med 2010;18:60.

[11] O'Reilly D, Mahendran K, West A, Shirley P, Walsh M, Tai N. Opportunities for improvement in the management of patients who die from haemorrhage after trauma. Br J Surg 2013;100:749–55.

[12] Teixeira PG, Inaba K, Hadjizacharia P, Brown C, Salim A, Rhee P, et al. Preventable or potentially preventable mortality at a mature trauma centre. J Trauma 2007;63:1338–46.

[13] Gruen RL, Jurkovich GJ, McIntyre LK, Foy HM, Maier RV. Patterns of errors contributing to trauma mortality: lessons learned from 2594 deaths. Ann Surg 2006;244:371–80.

[14] Acosta JA, Yang JC, Winchell RJ, Simons RK, Fortlage DA, Hollingsworth-Fridlund P, et al. Lethal injuries and time to death in a level I trauma centre. J Am Coll Surg 1998;186:528–33.

[15] Mackenzie CF, Wang Y, Hu PF, Chen SY, Chen HH, Hagegeorge G, et al. Automated prediction of early blood transfusion and mortality in trauma patients. J Trauma Acute Care Surg 2014;76:1379–85.

[16] Mackenzie CF, Gao C, Hu PF, Anazodo AN, Chen H, Dinardo T, et al. Comparison of decision-assist and clinical judgment of experts for prediction of life saving interventions. Shock 2014. In Press.

[17] Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37–46.

[18] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

[19] Kim YJ, Xiao Y, Mackenzie CF, Gardner SD. Availability of trauma specialists in level I and II trauma centers: a national survey. J Trauma 2007;63:676–83.

[20] Kwan I, Bunn F, Roberts I, WHO Pre-Hospital Trauma Care Steering Committee. Timing and volume of fluid administration for patients with bleeding. Cochrane Database Syst Rev 2003;3:CD002245.

[21] Hampton DA, Fabricant LJ, Differding J, Diggs B, Underwood S, De La Cruz D, et al. Prehospital intravenous fluid is associated with increased survival in trauma patients. J Trauma Acute Care Surg 2013;75:S9–15.

[22] Spinella PC, Holcomb JB. Resuscitation and transfusion principles for traumatic hemorrhagic shock. Blood Rev 2009;23:231–40.

[23] Barrera R, Nygard S, Sogoloff H, Groeger J, Wilson R. Accuracy of predictions of survival at admission to the intensive care unit. J Crit Care 2001;16:32–5.

[24] Birch K. Who benefits from Intensive Care in the field? J R Army Med Corps 2009;155:122–4.

[25] Detsky AS, Stricker SC, Mulley AG, Thibault GE. Prognosis, survival, and the expenditure of hospital resources for patients in an intensive-care unit. N Engl J Med 1981;305:667–72.

[26] Goettler CE, Waibel BH, Goodwin J, Watkins F, Toschlog EA, Sagraves SG, et al. Trauma intensive care unit survival: how good is an educated guess? J Trauma 2010;68:1279–87.

[27] Cloutier L, Ricci MA, Wald S, Eldrup-Jorgensen J, Bredenberg C. Predicting survival from ruptured abdominal aortic aneurysm. Computer modelling with AIM versus clinical judgment. Ann N Y Acad Sci 1996;800:234–5.

[28] Mackenzie CF, Hu P, Sen A, Dutton R, Seebode S, Floccare D, et al. Automatic pre-hospital vital signs waveform and trend data capture fills quality management, triage and outcome prediction gaps. AMIA Annu Symp Proc 2008;318–22.

[29] Dutton RP, Stansbury LG, Leone S, Kramer E, Hess JR, Scalea TM. Trauma mortality in mature trauma systems: are we doing better? An analysis of trauma mortality patterns, 1997–2008. J Trauma 2010;69:620–6.