OPEN ACCESS

Helicopter Emergency Medical Services attendance is associated with favourable survival outcomes in major trauma: derivation and internal validation of prediction models in a regional trauma system

Joanne Griggs ,[1,2] Jenny Harris,[2] Jack Barrett,[1,3] Scott Clarke,[1] John O'Neill,[4,5] Leigh Curtis,[6] Malcolm Tunnicliff,[4,7] Claire Baker,[8] Duncan Bootland,[9,10] Emma Ream,[2] Richard Lyon,[2,11] On behalf of Air Ambulance Charity Kent Surrey Sussex

Handling editor Richard Body

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/emermed-2025-215451).

For numbered affiliations see end of article.

Correspondence to
Joanne Griggs;
jogriggs@aakss.org.uk

Received 8 August 2025
Accepted 30 December 2025

To cite: Griggs J, Harris J, Barrett J, *et al*. *Emerg Med J* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/emermed-2025-215451

ABSTRACT

Background Survival benefit of Helicopter Emergency Medical Services (HEMS) attended major trauma remains inadequately quantified across injury severity. We evaluated HEMS performance and identified predictors of survival.

Methods Retrospective observational analysis of 3225 trauma patients attended by a regional HEMS in South-East England (2013–2022). Survival was assessed using W-statistic (W_s) methodology stratified by probability of survival (P_s) bands. Multivariable logistic regression identified predictors of 30-day mortality in major trauma (injury severity score (ISS) ≥15). Sub-analysis examined unexpected survival predictors and return of spontaneous circulation (ROSC) rates in traumatic cardiac arrest (TCA).

Results Among 2125 patients meeting W_s analysis criteria, observed (O) 30-day survival exceeded expected (E) survival (84.7% vs 81.3%; O/E ratio 1.04), yielding adjusted W_s of 5.23 (95% CI 3.27 to 7.19), representing 5.23 excess survivors per 100 patients. Survival benefit was greatest in severely injured patients with moderate survival probability (P_s 25–45%: 3.33 excess survivors per 100, 95% CI 1.37 to 5.29). Among patients with low probability of survival (P_s<50), 38.7% survived unexpectedly; younger ages and higher presenting Glasgow Coma Scale scores were key predictors of unexpected survival. Pre-hospital emergency anaesthesia (PHEA) was independently associated with unexpected survival in this group (adjusted OR 2.01, 95% CI 1.12 to 3.72, p=0.023). TCA ROSC rates demonstrated an annual improvement (6.3% increased odds per year, 95% CI 1.02 to 1.10, p=0.002).

Conclusion HEMS attendance to major trauma in this regional service was associated with survival exceeding case-mix adjusted predictions, and was most pronounced in severely injured patients. PHEA was associated with survival benefit in low probability patients, supporting the value of advanced pre-hospital interventions.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ International evidence for Helicopter Emergency Medical Services (HEMS) effectiveness in major trauma remains inconsistent due to methodological heterogeneity, inadequate sample sizes and variable outcome definitions. Previous studies have shown conflicting results regarding survival benefits, with limited granular analysis across injury severity bands to identify which patient populations derive the greatest benefit from this resource-intensive service.

INTRODUCTION

Trauma is a leading cause of death and disability worldwide, with optimal outcomes dependent on advanced interventions and expedient transfer to specialised trauma care.[1–3] Helicopter Emergency Medical Services (HEMS) represent a crucial component of many trauma systems, providing advanced clinical capabilities and expedited transport.[2 4 5] Operational demand on HEMS has increased over recent decades, raising questions regarding which trauma patient populations derive the most significant benefit from this resource-intensive, finite service.[6–8]

Traditional trauma system performance has been evaluated by comparing the ratio of observed (O) outcomes against predicted (E) outcomes, which scoring systems refer to as *expected* outcomes (O/E).[9] Performance is then measured using the W-statistic (W_s) which represents the number of excess survivors per 100 patients.[10] However, stratification of performance across different probability of survival bands (P_s) provides a more granular understanding of which patients trauma systems most benefit.[11] Of particular interest are patients who survive despite a low predicted probability of survival (either P_s<30% or P_s<50%), often termed 'unexpected survivors' (hereafter termed P_s<30 and P_s<50).[12] Unexpected survivors may offer insights into modifiable factors that contribute to positive mortality outcomes.[13 14]

Mortality predictors in trauma populations are well-established,[15 16] but conversely, few studies have identified patient and intervention factors associated with unexpected survival in severely injured patients (those with an Injury Severity Score (ISS) ≥15).[1 15 16] Identifying mortality predictors in HEMS-attended cohorts validates performance measurement methodology and provides clinical context for interpreting outcomes.[17] Further, these factors could enhance trauma outcome

WHAT THIS STUDY ADDS

⇒ This large regional cohort study demonstrates that in one UK HEMS, observed survival exceeded case-mix adjusted predictions by 5.23 per 100 major trauma patients, with the greatest effect observed in severely injured patients with moderate survival probability (25–45% P_s band). The study provides robust evidence using adjusted W-statistic methodology and identifies age, Glasgow Coma Scale score and pre-hospital interventions (such as PHEA) as key predictors of survival, while demonstrating significant annual improvement in TCA ROSC rates.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ These findings provide supportive evidence for continued investment in HEMS, particularly for severely injured patients, though comparative studies with alternative care pathways are needed to establish causal effectiveness. The validated prediction models may encourage real-time performance benchmarking and quality improvement initiatives, while the methodology provides a framework for trauma system evaluation that could inform international HEMS service development and clinical governance. We advocate for collaborative UK-wide HEMS data collection using standardised methodology to enable pooled analysis, to examine rare outcomes with adequate statistical power and establish national performance benchmarks.

benchmarking and performance measurement frameworks, particularly as evidence for clinical effectiveness of advanced trauma interventions delivered by specialist teams remains inconsistent.[18 19]

We aimed to: (1) evaluate regional trauma system performance by comparing observed-to-expected survival across probability bands; (2) identify independent predictors of 30-day mortality in HEMS-attended trauma sub-grouped by ISS ≥15; and (3) determine factors associated with unexpected survival in patients with low predicted survival probability.

METHODS

Study design and setting

A retrospective observational cohort study of all trauma patients attended and conveyed by the Air Ambulance Charity Kent Surrey Sussex (KSS) between 1 January 2013 and 31 December 2022. Patients were included if KSS clinicians attended and conveyed patients with a documented traumatic mechanism of injury. For survival analysis and performance benchmarking, patients were eligible if they met Trauma Audit and Research Network (TARN) inclusion criteria: hospital length of stay (LOS) >72 hours, death in hospital from injury, critical care admission or inter-hospital transfer for specialist care.[20] Patients pronounced life extinct (PLE) at the scene were excluded from primary survival analyses as they do not receive TARN P_s scores, but were included in a traumatic cardiac arrest (TCA) sub-analysis. Patients who died at the scene but underwent post-mortem examination were not included as TARN does not calculate P_s scores for pre-hospital deaths. Transparent Reporting of a multivariable Prediction Model for Individual Prognosis of Diagnosis (TRIPOD+AI 2024)[21] and the Reporting of studies Conducted using Observational Routinely Collected Health Data (RECORD) Statement were adhered to.[22]

Major trauma patients attended by KSS are conveyed to one of three major trauma centres (MTCs) in the South-East region. These include Royal Sussex County Hospital, Brighton (Sussex Trauma Network), St George's Hospital, London (South-West London and Surrey Trauma Network), or King's College Hospital, London (South-East London, Kent and Medway Trauma Network), covering a region of 7500 km^2 with a mixed rural–urban resident population of 4.5 million. Conveyance to an MTC may involve bypassing a trauma unit (TU). KSS operates 24/7 with two Leonardo Augusta Westland 169 (AW169) helicopters and two rapid response vehicles used interchangeably depending on operational need. Operating this model of care aims to deliver a doctor–paramedic team to a patient within 30 mins of an emergency 999/112 call.

Advanced interventions and deployment

Doctor–paramedic teams complement ground emergency medical service (EMS) provision, including a specialist critical care paramedic (CCP) workforce, by delivering advanced trauma interventions at the scene and during transport. Interventions include pre-hospital emergency anaesthesia (PHEA) and post-PHEA infusion, advanced ventilation strategies, blood component transfusion, advanced vascular access and invasive arterial blood pressure monitoring, and surgical procedures to include lateral canthotomy, thoracostomy and resuscitative thoracotomy (RT). Regional pathways comprise direct to CT, direct transfer for neurological intervention and massive transfusion protocols. Clinicians have 24/7 consultant and operational decision support. From the point of injury, a 999/112 call is made to the emergency operations centre and coded using NHS pathways.[23] HEMS deployment is determined through joint decision-making between co-located non-clinical dispatchers and a CCP at the associated EMS. Each trauma call is interrogated, triaged and an appropriate asset tasked according to pre-defined criteria, ensuring clinical oversight of tasking decisions.[24]

Blood component administration follows a haemostatic resuscitation approach prioritising early and balanced blood product transfusion in patients with suspected major haemorrhage. The standard approach involves balanced transfusion ratios, with packed red blood cells (PRBCs) and plasma (either freeze-dried plasma (FDP) or fresh frozen plasma (FFP)) administered in a 1:1 ratio, with the initial component at the discretion of the attending clinicians.

Patient eligibility and selection

Major trauma patients were eligible for inclusion if attended by KSS between 1 January 2013 and 31 December 2022. Patients were included if attended by KSS clinicians, regardless of transport modality (helicopter or ground vehicle) and irrespective of whether advanced interventions were delivered. Study size was determined by the number of patients meeting the eligibility criteria during the study duration.

Data sources and data acquisition

Data were extracted from the Electronic Patient Clinical Record (EPCR) HEMSBase 3.0 (MedicOne Systems) using Zoho Analytics. Notable variables included: *baseline demographic descriptors*, including age and sex; *scene descriptors*: critical care paramedic, blunt/penetrating, mechanism of injury (MOI); *physiological parameters*: Glasgow Coma Scale (GCS) score, respiratory rate (RR), heart rate (HR), oxygen saturation (SpO_2), end-tidal carbon dioxide ($EtCO_2$), systolic blood pressure (SBP) and diastolic blood pressure (DBP); *injury type*: cardiac arrest

(CA), injured body region (head, neck, thorax, abdomen/pelvis and limb); *trauma interventions*: PHEA, blood component transfusion including PRBC, FFP, FDP and anticoagulant reversal with Beriplex. Others included: endotracheal intubation (ETI) placement in CA, supraglottic airway, front-of-neck access, surgical airway, oropharyngeal airway, nasopharyngeal airway, positive pressure ventilation, thoracostomy, intercostal chest drain, thoracotomy, pelvic binder, haemostatic dressing, blast dressing, compression dressing, suture, maxillofacial intervention; *patient disposition*: conveyance status and injury severity; *outcomes*: 30-day survival, length of stay (LOS), ISS, and P_s (defined in *Definitions and calculation of survival scores*). A post-calculated, standardised case-mix adjusted outcome score was included for each trauma patient from a TARN participating centre.[25 26]

Presenting physiology refers to the first recorded vital signs on HEMS patient contact, before advanced interventions where possible. Final physiology refers to the last recorded measurements before hospital handover. ISS and P_s scores were calculated by TARN using in-hospital data including CT imaging, operative findings and diagnostic codes submitted by receiving trauma centres according to standard methodology.[22 25 27] Patients attended by HEMS are subsequently matched to TARN records to obtain retrospectively calculated ISS and P_s values. This approach is standard for HEMS outcome studies as anatomical injury scoring requires complete diagnostic information which is only available after hospital assessment.[1]

Study data were secured anonymously within the Research and Data Capture (REDCap) Data Management System hosted at KSS. REDCap is a secure, web-based application designed to support data capture for research which holds the required safeguards for data security and privacy.[27]

Data validation and pre-processing

Before data extraction, the primary EPCR database underwent quality assurance processes to ensure data validity and completion, where individual records are checked for errors and logic flaws. Anonymised data were subsequently downloaded into Microsoft Excel (Microsoft Corp, Redmond, WA, USA, version 16.79.1) for pre-processing and coding. Data extraction was performed by one investigator (SC), with a random sample of records independently verified by a second investigator (JG) to ensure accuracy. Discrepancies were resolved through consensus review of source documentation. Automated validation checks flagged logical inconsistencies for manual review.

Definitions and calculation of survival scores

Survival outcome was defined as alive/dead status at 30 days post-injury. Unexpected survivors were defined as those surviving to 30 days despite low predicted survival probability (P_s<30 and P_s<50).

Probability of survival (P_s) represents the expected survival probability for a trauma patient based on their specific combination of demographic, physiological and injury characteristics compared with similar patients in the historical TARN database. P_s is calculated by TARN using the P_s17 model incorporating age, sex, GCS, physiological parameters at hospital arrival and anatomical injuries determined from CT imaging and clinical findings.[20 27] The TARN prediction model has been validated in UK trauma populations demonstrating excellent discrimination (AUC 0.90, 95% CI 0.89 to 0.90) and represents the national standard for risk-adjusted trauma outcome benchmarking.[20] A P_s score represents retrospective probability; a score of 70% implies that 70 out of every 100 patients with that profile survived historically.

Expected outcomes (E) for performance evaluation are derived by summing individual P_s scores across all patients ($E=\Sigma P_s$). For example, if 100 patients each have P_s=70%, the expected survivors are E=70. Observed outcomes (O) are the actual number of patients surviving to 30 days.

Trauma system performance uses both crude (*W*) and case-mix adjusted (W_s) survival statistics.[28–30] Both represent excess survivors per 100 patients compared with expected outcomes. The *W* score is calculated as:

$$W = (\text{observed survivors (O)} - \text{expected survivors (E)}) \times 100 \div \text{total patients}$$

The O/E ratio was calculated as:

$$O \div E$$

The W_s statistic applies a weighting system using the formula $W_s=(\sum(O-E)\times\text{Weight})/(\sum\text{Weight})\times 100$, where $\text{Weight}=P_s\times(1-P_s)$. This gives greater emphasis to patients with moderate to severe survival probabilities (~50%) and less emphasis to those with very high or very low survival probabilities. A positive *Ws* indicates more survivors than expected and a negative score indicates fewer than expected survivors. These metrics were subsequently stratified by probability of survival (%) bands (*95–100*, *90–95*, 80–90, 65–80, 45–65, 25–45 and 0–25) using right-inclusive boundary classification common to this database.

For TCA secondary analyses, return of spontaneous circulation (ROSC) was defined as sustained return of circulation (palpable pulse, measurable blood pressure) maintained until hospital handover. ROSC represents a pre-hospital performance metric partly attributable to HEMS interventions.

Primary and secondary outcomes

Primary outcome

The primary outcome was standardised mortality ratio using observed-to-expected (O/E) mortality rate (W_s) across predicted survival probability bands (P_s) in HEMS-attended major trauma.

Secondary outcome(s)

Thirty-day mortality and clinical outcomes were examined across clinically relevant subgroups:

1. 30-day mortality in all HEMS-attended major trauma
2. 30-day mortality in severe trauma (ISS ≥15)
3. Unexpected survival rates in low predicted survival subgroups (P_s<30 and P_s<50)
4. Hospital and intensive care unit (ICU) LOS in severe trauma (ISS ≥15)
5. ROSC and 30-day mortality in TCA.

Statistical analysis

Continuous variables are presented as mean±SD or median and IQR. Categorical variables are presented as frequencies and percentages (%). Baseline characteristics were compared using χ^2 or Fisher's exact test for categorical variables and independent t-tests or Mann-Whitney U test for continuous variables, as appropriate.

Missing data were handled using multiple imputation (MI) by chained equations (MICE) with predictive mean matching generating five imputed datasets pooled from 50 iterations according to Rubin's rules.[31 32] All regression analyses used pooled MI results. Convergence diagnostics are shown in online supplemental table S4 and figure S5. Complete case sensitivity analysis showed consistent effect estimates with <10% difference

for most predictors confirming robustness to imputation method (online supplemental table S6).

Three complementary analyses were performed: (1) Observed/expected survival analysis: Crude and case-mix adjusted W_s statistics were calculated for each P_s band. Multivariable prediction models with logistic regression identified independent predictors of 30-day mortality. Variable selection for multivariable models was theory-driven, based on established trauma literature and clinical plausibility. Candidate variables were identified a priori and included: patient factors (age, sex), injury characteristics (mechanism, anatomical region, suspected traumatic brain injury (TBI, suspected traumatic haemorrhage), presenting physiology (GCS, SBP, heart rate, $EtCO_2$), and pre-hospital interventions (PHEA, ETI, thoracostomy, blood component transfusion, pelvic binder). All candidate variables with univariate associations ($p<0.10$) or strong theoretical rationale were entered into multivariable models. Variables were retained in final models based upon independent significance ($p<0.05$) or identification as important confounders based on clinical knowledge. Multicollinearity was assessed with correlated blood pressure variables removed to avoid redundancy. Clinically plausible two-way interactions were tested using likelihood ratio tests, with significant interactions retained in the final model. The events-per-variable (EPV) ratio exceeded recommended thresholds at 28:1. Model performance was assessed using McFadden's pseudo-R^2 and Akaike Information Criterion (AIC).[3] Additional variable selection was performed using Least Absolute Shrinkage and Selection Operator (LASSO) regression with 10-fold cross-validation, selecting the optimal regularisation parameter (λ) using the one-standard-error rule to balance parsimony with predictive performance.[33] Internal validation was undertaken with a 70/30 random data split. The prognostic model aimed to identify independent predictors of 30-day mortality rather than optimise clinical prediction or triage. Performance assessment prioritised overall discrimination (area under the curve (AUC)) and calibration across the full probability range. Calibration was evaluated using calibration plots and the Hosmer-Lemeshow (H-L) test, with discrimination via AUC. For prognostic research identifying risk factors, accurate probability estimates across all risk strata and appropriate risk ranking are essential, rather than optimising binary classification or sensitivity at a fixed threshold.

Secondary analyses for LOS (hospital and ICU) were restricted to patients surviving to discharge to avoid survival bias. Due to right-skewed distributions, LOS data were log-transformed and analysed using a negative binomial regression. Results are presented as odds ratios (OR) with 95% confidence intervals (95% CI). Data were pre-processed in Microsoft Excel (Microsoft Corp, Redmond, WA, USA). Statistical significance was pre-determined at $p<0.05$. Statistical analysis and visualisations were performed using R version 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria).

Ethical considerations

Data were routinely collected and met Health Research Authority (HRA, UK) criteria for service evaluation. Research

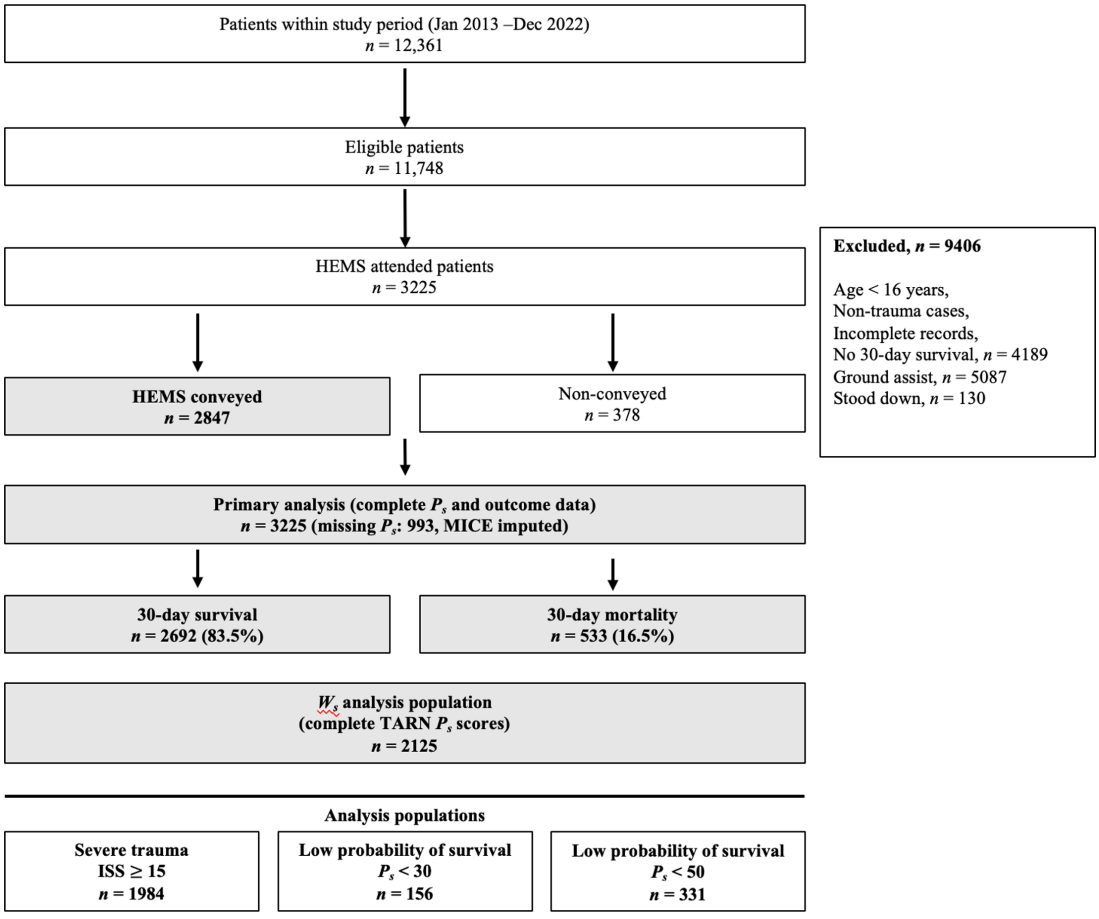

Figure 1 Flow diagram for derivation of study population stratified by survival to 30 days. Derivation of study population. GA, ground assist; HEMS, Helicopter Emergency Medical Services; ISS, Injury Severity Score; MICE, Multiple Imputed Chained Equations; PLE, pronounced life extinct; P_s, predicted survival score; TARN, Trauma Audit and Research Network.

ethics committee approval was not required, and the project was approved by the KSS Research and Innovation Committee under a TARN data sharing agreement. Major trauma systems provided letters of support for conducting this research. Study data were collected and managed using REDCap electronic data capture tools hosted securely with anonymised patient identifiers to ensure data protection and confidentiality.

RESULTS

Baseline demographics and patient characteristics

Of 12 361 patients screened, 3225 HEMS attended trauma patients with complete outcome data were included in the primary analysis, with 2125 eligible for W_s analysis. Additional subgroup analyses were conducted in patients with severe trauma (ISS ≥15, n=1984) and patients with low probability of survival (P_s<30, n=156; P_s<50, n=331) (figure 1). Median age was 43 years (IQR 25–61) with a predominance of males (74%). Blunt trauma accounted for 92% of cases, with road traffic collisions (58%) being the most common mechanism of injury. Median ISS was 24 (IQR 14–33), with 1984 patients (72.7%) classified as severe (ISS ≥15) and 605 (22.2%) as moderate (ISS 9–15). Patient demographics, injury characteristics, trauma interventions and clinical outcomes are stratified by survival or mortality at 30 days (table 1). Significant differences were observed across groups for age, mechanism of injury, ISS, presenting physiology, pre-hospital interventions (including PHEA, blood component transfusion and thoracostomy) and outcomes (p<0.001). Non-conveyed patients had significantly higher median P_s (99% vs 95.6%, p<0.001) validating on-scene transport decisions and our subsequent focus analysing conveyed patients (online supplemental table S1).

Observed and expected survival of HEMS attended patients

Among 2125 patients with complete P_s and outcome data, overall 30-day survival rate was 84.7% compared with expected 81.3%, yielding an adjusted W_s of 5.23 (95% CI 3.27 to 7.19, p<0.001) and O/E ratio of 1.04 (figure 2). This overall W_s exceeds individual band values due to variance-weighting methodology ($P_s \times (1-P_s)$), which gives greater influence on moderate survival probability bands (45–80%).

System performance showed positive values across all P_s bands with incremental positive values maintained throughout the study period (annual change: −0.138, 95% CI −0.40 to –0.13, p=0.26). Survival advantage was greatest in patients with moderate to severe injuries (P_s 25–45%, W_s 3.33, 95% CI 1.37 to 5.29, p<0.001), representing a 35% relative increase (49.1% actual vs 36.4% expected). A calibration plot is available in online supplemental figure S2.

Predictors of 30-day mortality

Multivariable logistic regression identified key independent predictors of 30-day mortality (table 2). Key predictors included: age (adjusted OR (aOR) 1.02 per year, 95% CI 1.01 to 1.04, p<0.001), GCS (aOR 0.63 per point, 95% CI 0.57 to 0.69, p<0.001), and pre-hospital intubation (aOR 3.66, 95% CI 1.91 to 6.99, p<0.001). Cardiac arrest (aOR 2.49, 95% CI 1.53 to 4.08, p<0.001) and traumatic haemorrhage (aOR 1.64, 95% CI 1.12 to 2.40, p=0.010) were associated with increased mortality, while penetrating injury was independently associated with reduced mortality (aOR 0.33, 95% CI 0.11 to 0.87, p=0.037). Thoracic injury demonstrated increased mortality risk in the adjusted model (aOR 2.75, 95% CI 1.27 to 5.95, p=0.010). A significant age×GCS interaction improved model fit (aOR 1.00, 95% CI 1.00 to 1.01, p<0.001). PHEA was associated with reduced mortality (aOR 0.55, p=0.021). The final model achieved excellent discrimination (AUC 0.919, 95% CI 0.907 to 0.930) with adequate calibration (H-L p=0.151).

Missing data patterns and imputation procedures are detailed in online supplemental figure S3 and S5, online supplemental table S4. Sensitivity analysis comparing complete case analysis with MI showed consistent effect estimates, with ORs differing by <10% (online supplemental table S6). LASSO regression validated core prognostic factors, with a parsimonious 7-variable model maintaining excellent performance (AUC 0.911, ΔAUC=0.007 vs full model) (online supplemental table S7).

Internal model validation

Internal model validation (70:30 training: test data split) was performed with stratified sampling to maintain mortality rates, with the final model trained on 2258 (70%) and validated on 967 (30%) patients. The model demonstrated excellent discrimination with a test AUC of 0.922 (95% CI: 0.90-0.94) and no evidence of overfitting (AUC difference: -0.007). Test set performance showed 85.4% accuracy, 88.5% sensitivity, 97.1% specificity, negative predictive value (NPV). The H-L test indicated acceptable calibration (χ^2 =15.9, p = 0.043) (Supplementary Figure S8)

Severely injured (ISS ≥15) subset analysis

Among 1984 HEMS-conveyed patients (ISS ≥15) 30-day mortality was 20.0% (n=397). In multivariable analysis, independent predictors were age (aOR 1.02 per year, p<0.001), presenting GCS (aOR 0.61 per point decrease, p<0.001), blood component transfusion (aOR 1.60, p=0.021), cardiac arrest (aOR 1.85, p=0.020) and ETI (aOR 2.80, 95% CI 1.40 to 5.60, p=0.004 (table 3). Abdomen-pelvic injury was reported (aOR 0.79, p=0.193) (table 3). An age×GCS interaction significantly improved prediction (ΔAIC=−21.6, p<0.001) with age effect stronger in patients with higher GCS (9.6% increased risk per year at GCS 15 compared with 3.2% at GCS 3). The model achieved strong discrimination (McFadden R^2=0.389).

Length of hospital and ICU stay in HEMS attended and ISS ≥15

Among 2075 patients, median hospital LOS was 15 days (IQR 8–29), and ICU stay was 2 days (IQR 0–8), increasing to 19 days (IQR 10–36) and 4 days (IQR 0–12), respectively, in severely injured patients (ISS ≥15, n=1397). Multivariable regression identified consistent predictors of longer hospital stay: higher injury severity (incidence rate ratio (IRR) 1.02–1.04) and reduced GCS (IRR 0.87–0.96), at p<0.001. Pre-hospital interventions reflecting disease severity were associated with increased LOS, including PHEA (hospital IRR 1.41, ICU IRR 2.63) and blood component transfusion (hospital IRR 1.44, ICU IRR 1.96), at p<0.001. Penetrating injury was associated with shorter stays (hospital IRR 0.60, ICU IRR 0.52, p<0.001).

Analysis of low probability of survival in patients with P_s<30 and P_s<50

Among 156 patients with P_s<30, 41 (26.3%) unexpectedly survived to 30 days, while among 331 patients with P_s<50, 128 (38.7%) survived unexpectedly (figure 3, online supplemental table S9). In adjusted analyses controlling for age and presenting GCS, younger age and higher GCS remained the primary independent predictors of unexpected survival (P_s<50: age aOR 0.98 per year, 95% CI 0.97 to 0.99, p<0.001; GCS aOR 1.19 per

Table 1 Patient demographics, injury characteristics and trauma interventions stratified by 30-day survival

Variable	Overall N=3225	30-day survival n=2692	30-day mortality n=533	P value
Baseline demographics				
Age (median (IQR)) (N=3225)	43 (25–61)	40 (25–57)	63 (39–78)	<0.001
Sex, male (n, %) (N=3225)	2434 (75.5)	2075 (77)	359 (67.3)	<0.001
Sex, female (n, %) (N=3225)	790 (24.5)	616 (22.9)	174 (32.6)	
Injury characteristics				
ISS (median (IQR)) (N=2729)	24 (14–33)	20 (13–29)	30 (25–41)	<0.001
Missing (n)	496	389	107	
Penetrating mechanism (n, %) (N=3225)	266 (8.2)	260 (9.7)	6 (1.1)	<0.001
Anatomical injury site				
Head injury (n, %) (N=3225)	1839 (57.0)	1401 (52.0)	438 (82.2)	<0.001
Neck injury (n, %) (N=3225)	284 (8.8)	215 (8.0)	69 (12.9)	<0.001
Thorax injury (n, %) (N=3225)	1281 (39.7)	1067 (39.6)	214 (40.2)	0.863
Abdomen/pelvis injury (n, %) (N=3225)	1179 (36.6)	1044 (38.8)	135 (25.3)	<0.001
Limb injury (n, %) (N=3225)	1398 (43.3)	1235 (45.9)	163 (30.6)	<0.001
Presenting physiology				
RR (median (IQR)) (N=2989)	17 (8–24)	18 (10–24)	15 (0–23)	<0.001
Missing (n)	237	217	20	
SpO_2 (median (IQR)) (N=3084)	97 (90–100)	98 (92–100)	93 (81–99)	<0.001
Missing (n)	141	122	19	
HR, bpm (median (IQR)) (N=3096)	86 (67–107)	86 (68–106)	86 (60–113)	0.567
Missing (n)	129	114	15	
SBP, mmHg (median (IQR)) (N=3059)	128 (107–146)	128 (109–145)	125 (90–154)	0.174
Missing (n)	166	138	28	
DBP, mmHg (median (IQR)) (N=3059)	80 (66–93)	81 (68–93)	78 (58–95)	0.002
Missing (n)	166	138	28	
MAP, mmHg (median (IQR)) (N=3059)	97 (81–110)	97 (83–110)	95 (70–113)	0.031
Missing (n)	166	138	28	
SI (median (IQR)) (N=2918)	0.70 (0.50–0.90)	0.70 (0.50–0.90)	0.70 (0.50–1.10)	0.058
Missing (n)	307	247	60	
Lactate (median (IQR)) (N=364)	2.7 (1.9–4.0)	2.7 (1.8–3.8)	3.2 (2.1–5.1)	0.017
Missing (n)	2861	2384	477	
GCS (median (IQR)) (N=3169)	14 (8–15)	14 (11–15)	4 (3-9)	<0.001
Missing (n)	56	49	7	
Final physiology				
RR (median (IQR)) (N=2989)	16 (12–21)	17 (12–22)	14 (10–18)	<0.001
Missing (n)	236	218	18	
SpO_2 (median (IQR)) (N=3084)	98 (95–100)	99 (96–100)	97 (86–99)	<0.001
Missing (n)	141	122	19	
HR, bpm (median (IQR)) (N=3096)	86 (70–104)	86 (70–103)	88 (68–109)	0.228
Missing (n)	129	114	15	
SBP, mmHg (median (IQR)) (N=3059)	124 (108–141)	124 (110–141)	122 (99–145)	0.010
Missing (n)	166	138	28	
DBP, mmHg (median (IQR)) (N=3059)	80 (67–92)	80 (68–92)	78 (61–92)	0.003
Missing (n)	166	138	28	
MAP, mmHg (median (IQR)) (N=3059)	95 (82–108)	96 (83–108)	93 (75–110)	0.006
Missing (n)	166	138	28	
Cardiac arrest and subsequent resuscitation (n, %) (N=3225)	164 (5.1)	53 (2.0)	111 (20.8)	<0.001
Trauma Interventions				
ETI (n, %) (N=3225)	1244 (38.6)	796 (29.6)	448 (84.1)	<0.001
PHEA (n, %) (N=3225)	1095 (34.0)	744 (27.6)	351 (65.9)	<0.001
Thoracostomy (n, %) (N=3225)	459 (14.2)	259 (9.6)	200 (37.5)	<0.001
RT (n, %) (N=3225)	11 (0.3)	1 (0.0)	10 (1.9)	<0.001
PRBC (n, %)* (N=3225)	155 (4.8)	118 (4.4)	37 (6.9)	<0.001
Plasma (n, %)*† (N=3225)	218 (6.8)	158 (5.9)	37 (6.9)	<0.001
Morbidity and mortality				
Hospital LOS (median (IQR)) (N=2602)	13 (6–25)	15 (8–29)	3 (1–7)	<0.001
Missing (n)	623	523	100	

Continued

Table 1 Continued

Variable	Overall N=3225	30-day survival n=2692	30-day mortality n=533	P value
ICU LOS (median (IQR)) (N=2648)	2 (0–8)	2 (0–8)	2 (1–5)	0.117
Missing (n)	577	477	100	
P_s (median (IQR)) (N=2232)	95.6 (74.0–99.0)	97.2 (86.9–99.2)	42.1 (22.2–63.2)	<0.001
Missing (n)	993	789	204	

Data presented as median (IQR) for continuous variables and n (%) for categorical variables. Statistical tests: Kruskal-Wallis test for continuous variables, χ^2 or Fisher's exact test for categorical variables. n denotes the number of patients with available data for that specific variable.
*Blood component presented as binary n (%) of those receiving any units.
†Freeze-dried plasma or fresh frozen plasma.
bpm, beats/min; DBP, diastolic blood pressure; ETI, endotracheal intubation; GCS, Glasgow Coma Scale; HR, heart rate; ICU, intensive care unit; LOS, length of stay; MAP, mean arterial pressure; PHEA, pre-hospital emergency anaesthesia; PRBC, packed red blood cells; P_s, predicted survival score; RR, respiratory rate; RT, resuscitative thoracotomy; SBP, systolic blood pressure; SI, shock index; SpO_2, saturations.

point, 95% CI 1.12 to 1.28, p<0.001; P_s<30: age aOR 0.98, 95% CI 0.96 to 0.99, p=0.014; GCS aOR 1.19, 95% CI 1.08 to 1.31, p<0.001).

Among intervention covariates, PHEA demonstrated a significant protective association with unexpected survival in P_s<50 patients (aOR 2.01, 95% CI 1.12 to 3.72, p=0.023). Blood transfusion (aOR 0.54, 95% CI 0.32 to 0.92, p=0.024) and thoracostomy (aOR 0.52, 95% CI 0.31 to 0.88, p=0.015) were inversely associated with survival, likely reflecting confounding by indication as these variables are preferentially performed in patients with severe physiological derangement. ETT and pelvic binder showed no significant independent associations. Sample size limitations in the P_s subgroup resulted in wide CIs for intervention estimates although thoracostomy (aOR 0.40, 95% CI 0.15 to 0.95, p=0.047) and pelvic binder (aOR 0.42, 95% CI 0.17 to 0.98, p=0.051) showed similar inverse patterns (online supplemental table S10).

Traumatic cardiac arrest and return of spontaneous circulation

Among 1316 patients presenting in TCA, 356 (27.1%) sustained ROSC to hospital and 960 patients were PLE at scene. Among the 356 ROSC patients, 30-day survival data were available for 185 (52%). Of these, 46 survived to 30 days (24.9% survival among ROSC with data, 12.9% of all ROSC patients). Post-ROSC in-hospital mortality was 75.1% (139/185). TCA ROSC rates demonstrated an annual improvement (6.3% increased odds per year, 95% CI 1.02 to 1.10, p=0.002).

Advanced trauma interventions were performed in 93.8% of TCA patients. In multivariable logistic regression adjusting for sex, age and pre-hospital interventions (n=1309), male sex was associated with reduced ROSC (aOR 0.62, 95% CI 0.38 to 1.00, p=0.049). ETI (aOR 7.05, 95% CI 3.23 to 15.38) and plasma units transfused (aOR 1.63 per unit, 95% CI 1.29 to 2.07)

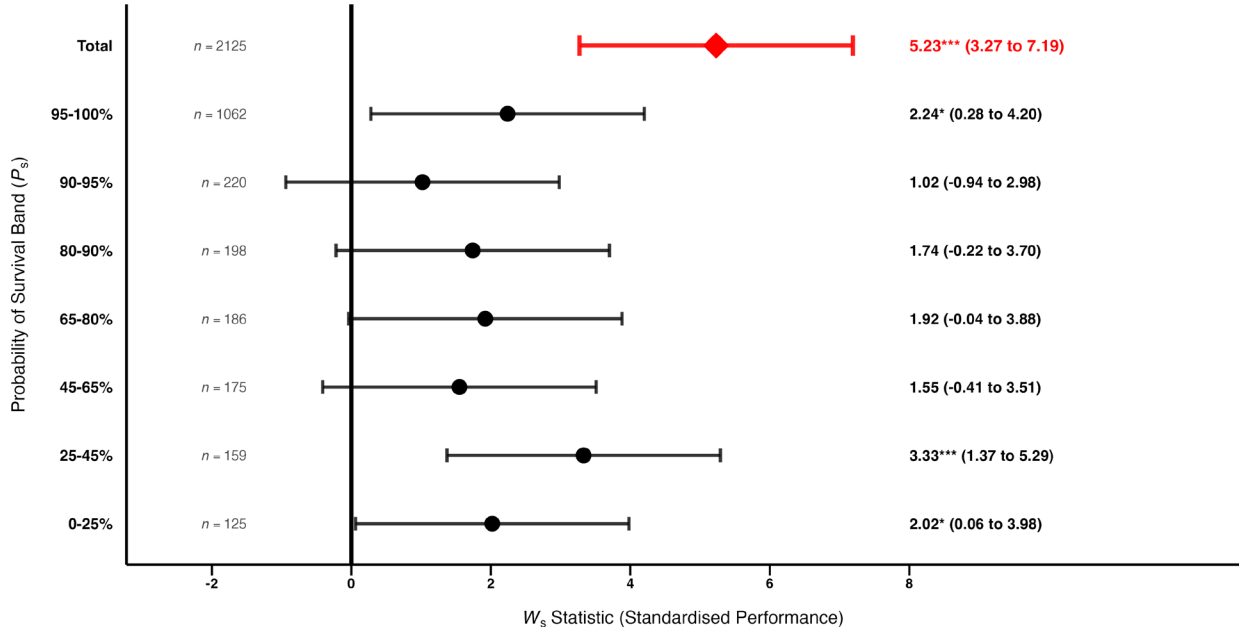

Figure 2 Observed and expected trauma survival across predicted survival categorisation (n=2125). W_s Statistic; adjusted W-statistic accounting for case-mix, calculated as (Observed−Expected)/√Variance. Positive values indicate better than expected performance with a reference line (W_s=0) indicating expected performance based on case-mix adjusted predictions, and error bars showing 95% CI. Black circles represent individual probability of survival bands; red diamond represents overall performance. Sample sizes shown on left (n=patients per band). Overall survival: 84.7% observed versus 81.3% expected (O/E ratio=1.04). Greatest benefit in 25–45% band: 78 actual versus 58 expected survivors (W_s=3.33). Overall W_s (5.23) exceeds individual effects due to weighting that emphasises moderate P_s bands. Statistical significance denoted as *p<0.05, ***p<0.001. Analysis restricted to patients with complete P_s and 30-day outcome data. P_s, predicted survival score.

Table 2 Univariate and multivariable logistic regression analysis for predictors associated with 30-day mortality (n=3225)

	Univariate analysis		Multivariable analysis	
Variable	**OR (95% CI)**	**P value**	**aOR (95% CI)**	**P value**
Demographics				
Age (years)	1.03 (1.03 to 1.04)	<0.001	1.02 (1.01 to 1.04)	<0.001
Sex (male)	0.61 (0.50 to 0.75)	<0.001	0.88 (0.67 to 1.16)	0.373
Physiological variables				
GCS	0.73 (0.72 to 0.75)	<0.001	0.63 (0.57 to 0.69)	<0.001
Presenting SBP (mmHg)	1.00 (1.00 to 1.00)	0.149	1.00 (1.00 to 1.00)	0.702
Heart rate	1.00 (1.00 to 1.00)	0.291	1.00 (1.00 to 1.00)	0.292
$EtCO_2$ (kPa)	0.95 (0.91 to 0.99)	0.012	0.95 (0.90 to 1.00)	0.072
Injury characteristics				
TBI	4.25 (3.38 to 5.40)	<0.001	1.05 (0.75 to 1.48)	0.777
Thoracic injury	1.02 (0.84 to 1.23)	0.825	2.75 (1.27 to 5.95)	0.010
Abdomen-pelvic injury	0.54 (0.43 to 0.66)	<0.001	0.85 (0.61 to 1.19)	0.351
Penetrating injury	0.11 (0.04 to 0.22)	<0.001	0.33 (0.11 to 0.87)	0.037
Cardiac arrest	13.10 (9.34 to 18.58)	<0.001	2.49 (1.53 to 4.08)	<0.001
Intervention variables				
PHEA	5.05 (4.15 to 6.16)	<0.001	0.55 (0.33 to 0.91)	0.021
ETI	12.55 (9.87 to 16.15)	<0.001	3.66 (1.91 to 6.99)	<0.001
Thoracostomy	5.64 (4.54 to 7.01)	<0.001	1.25 (0.81 to 1.94)	0.310
Pelvic binder	0.99 (0.82 to 1.19)	0.910	1.03 (0.78 to 1.36)	0.847
Blood component transfusion	3.40 (2.75 to 4.19)	<0.001	1.64 (1.12 to 2.40)	0.010
Interactions				
Age (years)×GCS	–	–	1.00 (1.00 to 1.01)	<0.001

Multivariable analysis of 30-day mortality risk factors.
Results from multiple imputation (m=5).
aOR, adjusted OR; $EtCO_2$, end-tidal carbon dioxide; ETI, endotracheal intubation; GCS, Glasgow Coma Scale; PHEA, pre-hospital emergency anaesthesia; SBP, systolic blood pressure; TBI, traumatic brain injury.

Table 3 Univariate and multivariable logistic regression analysis for mortality risk factors in severe trauma (ISS ≥15), n=1984

	Univariate analysis		Multivariable analysis*	
Variable	**OR (95% CI)**	**P value**	**aOR (95% CI)**	**P value**
Demographics				
Age (years)	1.03 (1.03 to 1.04)	<0.001	1.02 (1.01 to 1.04)	<0.001
Sex (male)	0.63 (0.51 to 0.78)	<0.001	0.86 (0.64 to 1.15)	0.302
Physiological variables				
Presenting GCS	0.75 (0.73 to 0.77)	<0.001	0.61 (0.55 to 0.68)	<0.001
Presenting SBP (mmHg)	1.00 (1.00 to 1.00)	0.443	1.00 (1.00 to 1.00)	0.224
Heart rate	1.00 (1.00 to 1.00)	0.379	1.00 (1.00 to 1.01)	0.162
Presenting $EtCO_2$ (kPa)	0.96 (0.92 to 1.00)	0.053	0.95 (0.90 to 1.01)	0.105
Injury characteristics				
TBI	3.00 (2.34 to 3.88)	<0.001	0.90 (0.62 to 1.31)	0.593
Thoracic injury	0.89 (0.73 to 1.09)	0.263	1.01 (0.71 to 1.45)	0.947
Abdomen-pelvic injury	0.53 (0.42 to 0.66)	<0.001	0.79 (0.55 to 1.12)	0.193
Penetrating injury	0.21 (0.07 to 0.47)	<0.001	0.57 (0.15 to 1.77)	0.363
Cardiac arrest	8.82 (6.15 to 12.82)	<0.001	1.85 (1.11 to 3.12)	0.020
Intervention variables				
PHEA	3.80 (3.08 to 4.70)	<0.001	0.68 (0.40 to 1.16)	0.163
ETI	9.03 (6.98 to 11.83)	<0.001	2.80 (1.40 to 5.60)	0.004
Thoracostomy	4.05 (3.22 to 5.08)	<0.001	1.34 (0.85 to 2.10)	0.207
Pelvic binder	0.83 (0.68 to 1.01)	0.070	0.96 (0.72 to 1.29)	0.788
Blood component transfusion†	2.54 (2.03 to 3.17)	<0.001	1.60 (1.07 to 2.37)	0.021
Interactions				
Age (years)×GCS	–	–	1.00 (1.00 to 1.01)	<0.001

Multivariable model including candidate variables + (Age×GCS) interaction. n=1984 conveyed patients with ISS ≥15, mortality rate 397 (20.0%).
*Model performance R^2 = 0.389, ΔAIC = 21.6.
†Blood component transfusion treated as binary variable. Results from multiple imputation (m=5).
AIC, Akaike Information Criterion; aOR, adjusted OR; $EtCO_2$, end-tidal carbon dioxide; ETI, endotracheal intubation; GCS, Glasgow Coma Scale; ISS, Injury Severity Score; PHEA, pre-hospital emergency anaesthesia; SBP, systolic blood pressure; TBI, traumatic brain injury.

were independently associated with achieving ROSC. Pelvic binder application was also associated with ROSC (aOR 1.91, 95% CI 1.11 to 3.28, p=0.019). Thoracostomy was negatively associated with ROSC (aOR 0.13, 95% CI 0.07 to 0.25), likely reflecting confounding by indication. Model discrimination was AUC=0.783 (online supplemental table S11).

DISCUSSION

This study sought to evaluate the performance of a regional trauma system by analysing observed versus expected survival rates across different probability bands. Our findings indicate that observed survival significantly exceeded case-mix adjusted predictions among 2125 HEMS attended trauma patients, with 5.23 excess survivors per 100 patients (95% CI 3.27 to 7.19). Positive system performance was maintained throughout the study period. Differential performance was observed across injury severity. The 35% relative survival increase in the 25–45% probability band (49.1% actual vs 36.4% expected) suggests that the observed survival advantage over predictions is greatest in patients with moderate to severe injuries.

While the adjusted W_s statistic of 5.23 indicates excess survivors per 100 patients compared with expected outcomes, it is important to interpret this in clinical context. Based on the annual volume of HEMS attended trauma patients in our system, this effect size may translate to approximately 115 additional lives saved per year. This estimate should be interpreted cautiously, as it represents excess survival compared with model predictions rather than a demonstrated causal effect of HEMS attendance and assumes consistent performance and case-mix over time. Nevertheless, it illustrates the potential magnitude of clinical benefit, consistent with previous economic and social benefits demonstrated in previous studies.[34 35] The precision of our findings is supported by narrow CIs (95% CI 3.27 to 7.19) for the W_s estimate.

International studies present conflicting evidence regarding HEMS effectiveness, likely reflecting methodological heterogeneity including inadequate sample sizes and inconsistent outcome definitions rather than true differences in clinical effectiveness.[7 36–39] While some studies report minimal survival advantages,[36 37] others demonstrate substantial benefits in heterogenous trauma populations.[1 3 17 35 40–42 43] Notably, studies showing differential effectiveness across injury severity align with our findings. One UK study found minimal overall benefit (0.9 excess survivors per 100 patients, p=0.58) but meaningful improvement in major trauma subgroups (4.5 excess survivors per 100, p=0.03).[2] Northern England demonstrates remarkably similar results to ours, with an enhanced care team W_s of 3.22 versus a non-enhanced care W_s of −2.97, yielding a difference of 6.18 (95% CI 3.19 to 9.17),[1] supporting our W_s methodology and the magnitude of effect observed.

Derivation of prognostic models demonstrates excellent performance for case-mix adjustment. The primary model achieved robust discrimination (AUC 0.919, 95% CI 0.907 to 0.930) with acceptable calibration (H-L p=0.151), and internal validation confirmed no evidence of overfitting (training AUC

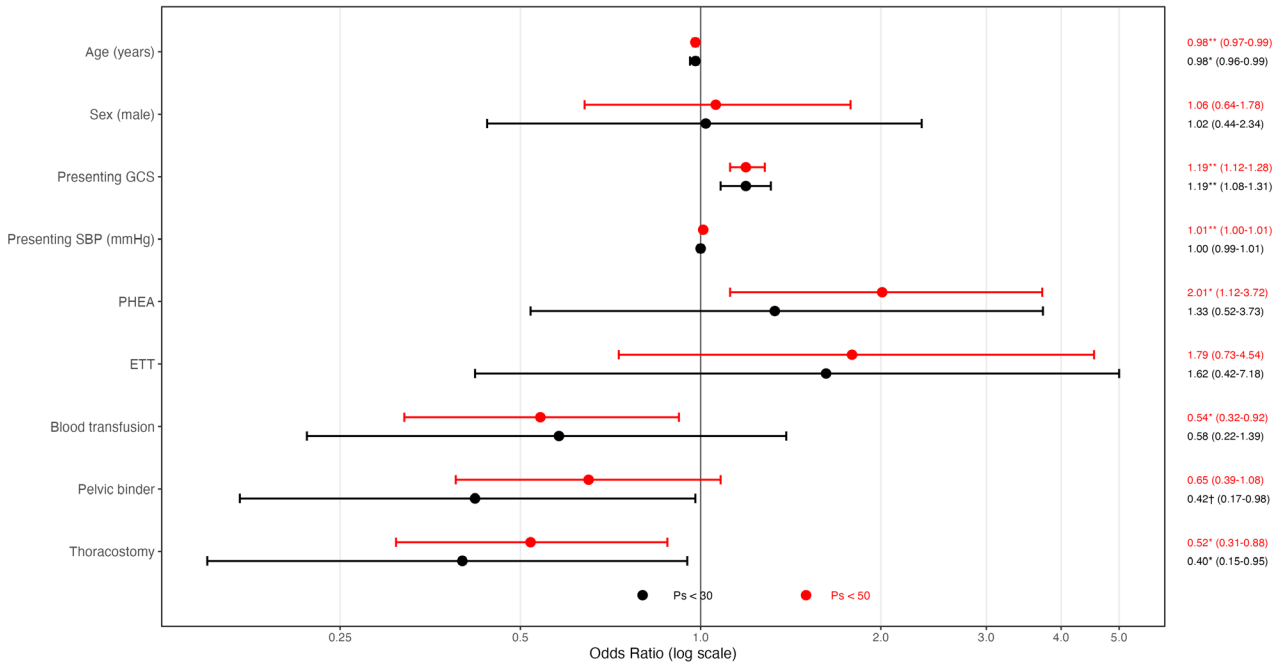

Figure 3 Forest plot showing prognostic factors for unexpected survival in both the P_s<30 (n=156, 41 survivors) and P_s<50 (n=331, 128 survivors) cohorts. Each intervention variable was adjusted separately for age and presenting GCS; age and GCS were mutually adjusted. Confidence intervals exceeding the display range are truncated at OR=5. ETT, endotracheal intubation; GCS, Glasgow Coma Scale; PHEA, pre-hospital emergency anaesthesia; P_s, predicted survival score. SBP, systolic blood pressure. Age: Consistent protective factor with younger age (adjusted OR (aOR) 0.98 in both groups). GCS: Strongest predictor (aOR 1.19*** in both groups). PHEA: Significant protective effect in P_s<50 (aOR 2.01*, p=0.023). Blood transfusion and thoracostomy: Inverse associations (aOR <1) likely reflect confounding by indication as these interventions are performed in patients with more severe physiological derangement. Injury Severity Score (ISS) excluded to prevent circularity with P_s score. Statistical significance denoted by *p<0.05, **p<0.01, ***p<0.001, †p<0.1.

0.915, test AUC 0.922). Age, GCS and ETI emerged as the strongest independent predictors of 30-day mortality. The significant age×GCS interaction reveals that older patients with preserved neurological function face disproportionately higher mortality risk, highlighting the complex interplay between physiological reserve and injury severity. LASSO regression validated core prognostic factors, confirming that a parsimonious 7-variable model maintains excellent performance with minimal loss (ΔAUC=0.007).

Multivariable analysis in the full cohort shows PHEA associated with reduced mortality (aOR 0.55, 95% CI 0.33 to 0.91). This is consistent with the association observed in the low probability subgroup analysis (aOR 2.01) and likely reflects the clinical benefit of securing definitive airway management despite being performed preferentially in more severely injured patients. Similarly, penetrating injury was associated with reduced mortality (aOR 0.31, 95% CI 0.10 to 0.83), aligning with previous studies demonstrating better outcomes when rapid surgical intervention is available.[44 45 46] Alternatively, our trauma system purposefully prioritises the triage of central penetrating trauma to an MTC, which may explain the seemingly protective effect.[47 48]

Low probability of survival patients (P_s<30 and P_s<50) had unexpected survival of 38.7% and 26.3%, respectively. Younger age (aOR 0.97 per year, 95% CI 0.96 to 0.98) and higher GCS (aOR 1.22 per point, 95% CI 1.15 to 1.30 for P_s<50) emerged as key predictors of unexpected survival, underscoring the prognostic importance of physiological reserve and neurological status beyond what conventional prediction models capture. In contrast, anatomical injury site was not independently associated after adjustment, suggesting that other patient factors and physiological response may be more important determinants in patients with low probability of survival.[47] This finding challenges the traditional emphasis on anatomical scoring systems and is suggestive of a holistic physiological approach to trauma prognostication.

Expanded analysis of intervention associations in low probability patients revealed that PHEA was independently associated with unexpected survival in P_s<50 patients (aOR 2.01, 95% CI 1.12 to 3.72, p=0.023). This association persisted after adjustment for age and GCS, suggesting clinical benefit beyond confounding by severity and supporting the value of definitive airway management in severely injured patients. In contrast, blood transfusion and thoracostomy showed inverse associations with survival (aOR 0.54 and 0.52, respectively), likely reflecting confounding by indication, as they are performed in patients with severe haemorrhagic and thoracic injuries. The absence of significant associations for ETI and pelvic binder may reflect a lack of independent effect, insufficient statistical power or complex confounding. Propensity-matched or instrumental variable approaches would be required to isolate true intervention effects, but were not feasible given sample size constraints. Intervention effects were clearer in the TCA subpopulation.

Among TCA patients, 356 (27.1%) achieved ROSC and were conveyed to hospital. Among ROSC patients, 30-day survival data were available for 185 (52%) with 46 surviving (24.9% survival among ROSC with data, 12.9% of all ROSC patients). Post-ROSC in-hospital mortality was 75.1%. ROSC rates demonstrated annual improvement (6.3% increased odds per year), increasing from 12.5% in 2013 to 27.4% in 2022. These results demonstrate meaningful survival in selected TCA patients. In

multivariable analysis (n=1309), male sex was associated with reduced ROSC (aOR 0.62, 95% CI 0.38 to 1.00, p=0.049), warranting further investigation. ETI (aOR 7.05, 95% CI 3.23 to 15.38) and plasma transfusion (aOR 1.63 per unit, 95% CI 1.29 to 2.07) were independently associated with achieving ROSC. Pelvic binder application was also associated with ROSC (aOR 1.91, 95% CI 1.11 to 3.28, p=0.019). Thoracostomy was negatively associated with ROSC (aOR 0.13, 95% CI 0.07 to 0.25). The substantial post-ROSC mortality (75.1%) indicates that while ROSC is necessary for survival, continued definitive hospital care remains critical. However, given advances in early trauma practice, ROSC rates may represent an important metric for evaluating pre-hospital trauma system effectiveness.

Phase-specific performance metrics in pre-hospital trauma systems require further exploration. Current predicted survival scores rely on hospital arrival physiology which may already reflect HEMS intervention benefits, potentially confounding the assessment of pre-hospital effectiveness. This limitation is exemplified by GCS measurements; in HEMS cohorts, patients may receive PHEA, making hospital arrival GCS unreliable for baseline assessment.[49 25] Missing or altered physiological data may also reduce the precision of survival predictions and limit accurate performance evaluation. Nevertheless, international studies have effectively incorporated risk estimation tools to account for missing variables, evidencing robust methodology to develop phase-specific performance metrics.[50] We demonstrate that pre-hospital variables achieve excellent discrimination (AUC 0.919) supporting the feasibility of developing real-time performance metrics that could enable continuous benchmarking and quality improvement. Machine-learning approaches may provide automated, phase-specific performance measurement for operational, governance and strategic oversight.[45] Recurrent neural networks analysing continuous monitoring data offer the ability to transform static predictors into dynamic assessments and such systems may encourage a move from retrospective analysis to prospective modelling.

Our findings demonstrate that observed survival exceeded case-mix adjusted predictions in HEMS-attended major trauma in our region, particularly in moderately to severely injured patients (P_s 25–45%, W_s 3.33). The identified mortality predictors serve two key purposes. First, they validate our measurement methodology; the associations between age, GCS and pre-hospital interventions with outcomes demonstrate that our models achieve excellent discrimination (AUC 0.919) for case-mix adjustment, enabling fair comparison of observed versus expected survival across different patient populations and time periods. This supports robust benchmarking and detection of temporal trends in system performance. Second, while not directly mandating protocol changes, this knowledge contextualises decision-making and reinforces physiological principles underlying current trauma management. Further, these findings establish a foundation for sophisticated performance monitoring.

Limitations inherent to observational studies are evident. First, we demonstrate associations but cannot establish causality between interventions and outcomes. We adopted W-statistic methodology rather than propensity score matching because valid control groups are unavailable and HEMS is routinely dispatched to the most severely injured patients, creating systematic case-mix differences that cannot be adequately matched against ground EMS cohorts. Absence of a ground EMS comparison group prevents determining whether benefits reflect HEMS-specific interventions versus the broader regional trauma system. Our findings demonstrate survival exceeding model predictions but cannot establish a causal relationship, as model calibration imperfections or unmeasured confounders could contribute to observed differences. While a randomised controlled trial would provide definitive evidence, this is neither ethical nor feasible given established HEMS integration into trauma systems. The 10-year period encompasses changes in protocols and technology confounding temporal trends. Additionally, the P_s score uses hospital (post-intervention) arrival data, meaning expected survival already reflects HEMS effects; our observed 5.23 excess survivors per 100 may thus underestimate true benefit. Second, unmeasured confounders (comorbidities, frailty) may influence outcomes, and dispatch involves complex triage introducing potential selection bias. Blood component transfusion suffers confounding by indication, that is, plasma recipients likely represent more severe cases requiring escalation. Third, no a priori sample size was calculated, though the events-per-variable ratios exceeded traditional thresholds (>28:1). Recent guidance suggests larger samples optimise performance and external validation is needed. Our model's 88.5% sensitivity reflects class imbalance and is adequate for benchmarking, but would require recalibration for further clinical applications.

Regarding data completeness, many patients had incomplete data for W-statistic analysis, with attrition primarily through non-TARN submission and not meeting inclusion criteria. Importantly, we cannot definitively establish missing at random assumptions, and if data were missing not at random, our estimates may be biased. Fourth, we could not reliably distinguish from electronic patient records those patients in cardiac arrest on HEMS arrival. This limits interpretation of TCA interventions, as pre-arrest interventions (eg, PHEA, blood transfusion) may differ from peri-arrest interventions. Also, among TCA patients achieving ROSC, 75.1% died after hospital admission, indicating substantial post-resuscitation mortality. We lack data on ethnicity, neurological outcomes and quality of life, limiting our assessment of potential disparities in outcomes across ethnic groups and long-term patient-centred outcomes. Finally, findings from one regional service may not generalise to different pre-hospital configurations. Unexpected survival analysis used conventional P_s thresholds; machine learning or dynamic physiological assessment incorporating frailty and functional status may provide more comprehensive understanding.

Clinical and research implications

The findings validate current practices including low-threshold ETI and blood component resuscitation in TCA, while highlighting the need for enhanced monitoring in older trauma patients given the age×GCS interaction demonstrating disproportionate mortality risk. The pronounced survival association in moderately to severely injured patients (P_s 25–45%) supports prioritising HEMS resources to this severity band. Beyond local implications, we advocate for collaborative UK-wide HEMS data collection using standardised methodology. Regional studies are limited by sample size, single-service biases and the inability to distinguish service-specific from system-level effects. A coordinated national approach with agreed outcome measures, standardised definitions and data sharing infrastructure would enable pooled analysis and the examination of rare outcomes with adequate power to establish meaningful national benchmarks. This would accelerate evidence generation and inform service development.

CONCLUSION

HEMS attendance to major trauma in this regional trauma system was associated with survival exceeding case-mix adjusted

predictions (5.23 excess survivors per 100 patients, 95% CI 3.27 to 7.19), with greatest effect in severely injured patients with moderate survival probability (P_s 25–45%, W_s 3.33). PHEA was independantly associated with unexpected survival in P_s <50. Our findings provide supportive evidence for advanced pre-hospital trauma interventions and demonstrate that phase-specific benchmarking can effectively evaluate HEMS performance, though comparative studies are needed to establish causal effectiveness.

Author affiliations
[1]Research and Innovation, Air Ambulance Kent Surrey Sussex, Redhill, UK
[2]School of Health Sciences, University of Surrey, Guildford, UK
[3]Research Department, South East Coast Ambulance Service NHS Foundation Trust, Banstead, UK
[4]Clinical Governance, Air Ambulance Kent Surrey Sussex, Redhill, UK
[5]Emergency Medicine, Southampton University Hospitals NHS Trust, Southampton, UK
[6]Air Ambulance Kent Surrey Sussex, Redhill, UK
[7]Emergency Medicine, King's College Hospital, London, UK
[8]Dyson School of Design Engineering / Imperial-X, Imperial College London, London, UK
[9]Medical Director, Air Ambulance Kent Surrey Sussex, Redhill, UK
[10]Emergency Medicine, Royal Sussex County Hospital, Brighton, UK
[11]Executive Director of Research and Innovation, Air Ambulance Kent Surrey Sussex, Redhill, UK

Acknowledgements The authors gratefully acknowledge the invaluable contributions of the clinical, operational and senior management teams at Air Ambulance Charity Kent Surrey Sussex. We extend particular thanks to Sophy Dunn for her meticulous efforts in data entry and validation processes at KSS. We acknowledge the academic critique provided by colleagues within the Faculty of Health Sciences at the University of Surrey, whose expertise significantly enhanced this study. Special recognition is due to the clinicians, research staff and senior management teams at both South-East Coast Ambulance Trust and across major trauma systems in the South-East region, whose collaboration made this research possible. The authors declare that this work would not have been feasible without the collective expertise and commitment of all individuals and organisations.

Contributors All authors fulfilled the CRediT taxonomy for authorship. Conceptualisation: JG, JH, ER, RL. Methodology: JG, JH, JB, ER, RL. Data curation: JG, SC. Formal analysis: JG, JH. Writing – first draft: JG. Writing – review and editing: JG, JH, JB, JON, LC, DB, CB, ER, RL. All authors read and approved the final manuscript. JG is the guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Representatives from each major trauma system and the internal charity board expressed support for research into trauma outcomes. While patients were not directly involved in study design, recruitment, or conduct, the results and clinical interpretation will be shared with lay representatives and partner organisations within the regional trauma networks to ensure broad stakeholder engagement and knowledge translation.

Patient consent for publication Not applicable.

Ethics approval All data were routinely collected and met Health Research Authority (HRA, UK) criteria as service evaluation. Research Ethics Committee approval was not required. The project was approved by the KSS Research and Innovation Committee and registered with the University of Surrey. Study conduct was in accordance with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. Analytical code is available on reasonable request.

ORCID iD
Joanne Griggs https://orcid.org/0000-0003-3810-4342

REFERENCES

1 Smith CA, Hardern RD, LeClerc S, *et al*. Prehospital analysis of northern trauma outcome measures: the PHANTOM study. *Emerg Med J* 2019;36:213–8.
2 Maddock A, Corfield AR, Donald MJ, *et al*. Prehospital critical care is associated with increased survival in adult trauma patients in Scotland. *Emerg Med J* 2020;37:141–5.
3 Lyons J, Gabbe BJ, Rawlinson D, *et al*. Impact of a physician-critical care practitioner pre-hospital service in Wales on trauma survival: a retrospective analysis of linked registry data. *Anaesthesia* 2021;76:1475–81.
4 Galvagno SMJ. Comparative effectiveness of helicopter emergency medical services compared to ground emergency medical services. *Crit Care* 2013;17:169.
5 Andruszkow H, Schweigkofler U, Lefering R, *et al*. Impact of Helicopter Emergency Medical Service in Traumatized Patients: Which Patient Benefits Most? *PLoS ONE* 2016;11:e0146897.
6 Butler DP, Anwar I, Willett K. Is it the H or the EMS in HEMS that has an impact on trauma patient mortality? A systematic review of the evidence. *Emerg Med J* 2010;27:692–701.
7 Hyde P, Mackenzie R, Ng G, *et al*. Availability and utilisation of physician-based pre-hospital critical care support to the NHS ambulance service in England, Wales and Northern Ireland. *Emerg Med J* 2012;29:177–81.
8 Brown JB, Gestring ML, Guyette FX, *et al*. Helicopter transport improves survival following injury in the absence of a time-saving advantage. *Surgery* 2016;159:947–59.
9 Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. Trauma Score and the Injury Severity Score. *J Trauma* 1987;27:370–8.
10 Hollis S, Yates DW, Woodford M, *et al*. Standardized comparison of performance indicators in trauma: a new approach to case-mix variation. *J Trauma* 1995;38:763–6.
11 Joosse P, Saltzherr T-P, van Lieshout WAM, *et al*. Impact of secondary transfer on patients with severe traumatic brain injury. *J Trauma Acute Care Surg* 2012;72:487–90.
12 Gruen DS, Guyette FX, Brown JB, *et al*. Association of Prehospital Plasma With Survival in Patients With Traumatic Brain Injury: A Secondary Analysis of the PAMPer Cluster Randomized Clinical Trial. *JAMA Netw Open* 2020;3:e2016869.
13 Hess JR, Brohi K, Dutton RP, *et al*. The Coagulopathy of Trauma: A Review of Mechanisms. *J Trauma* 2008;65:748–54.
14 Hemmila MR, Cain-Nielsen AH, Wahl WL, *et al*. Regional collaborative quality improvement for trauma reduces complications and costs. *J Trauma Acute Care Surg* 2015;78:78–85.
15 Kehoe A, Rennie S, Smith JE. Glasgow Coma Scale is unreliable for the prediction of severe head injury in elderly trauma patients. *Emerg Med J* 2015;32:613–5.
16 Den Hartog D, Romeo J, Ringburg AN, *et al*. Survival benefit of physician-staffed Helicopter Emergency Medical Services (HEMS) assistance for severely injured patients. *Injury* 2015;46:1281–6.
17 Thompson L, Hill M, Davies C, *et al*. Identifying pre-hospital factors associated with outcome for major trauma patients in a regional trauma network: an exploratory study. *Scand J Trauma Resusc Emerg Med* 2017;25:83.
18 Lockey DJ. Research questions in pre-hospital trauma care. *PLoS Med* 2017;14:e1002345.
19 Wilson MH, Habig K, Wright C, *et al*. Pre-hospital emergency medicine. *Lancet* 2015;386:2526–34.
20 Bouamra O, Jacques R, Edwards A, *et al*. Prediction modelling for trauma using comorbidity and 'true' 30-day outcome. *Emerg Med J* 2015;32:933–8.
21 Collins GS, Moons KGM, Dhiman P, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024;385:e078378.
22 Benchimol EI, Smeeth L, Guttmann A, *et al*. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12:e1001885.
23 Green J, Ewings S, Wortham R, *et al*. PP13 The NHS pathways (NHSP) medical call triage system and new 'nature of call' telephone screening tool, employed prior to NHSP, can accurately identify high acuity patients: analysis of sensitivity and specificity using routine ambulance service data. *Emerg Med J* 2019;36:e5.
24 Munro S, Joy M, de Coverly R, *et al*. A novel method of non-clinical dispatch is associated with a higher rate of critical Helicopter Emergency Medical Service intervention. *Scand J Trauma Resusc Emerg Med* 2018;26:84.

25 National Clinical Guideline Centre (UK). TARN: background and statistical techniques considered for analysis. In: *Major trauma: assessment and initial management*. National Institute for Health and Care Excellence (NICE), 2016. Available: https://www.ncbi.nlm.nih.gov/books/NBK368123/
26 NHS England Digital. Information standards for health and care. Available: https://digital.nhs.uk/data-and-information/information-standards [Accessed 16 Jun 2025].
27 Bouamra O, Wrotchford A, Hollis S, *et al*. A new approach to outcome prediction in trauma: a comparison with the TRISS model. *J Trauma* 2006;61:701–10.
28 Patridge EF, Bardyn TP. Research Electronic Data Capture (REDCap). *J Med Libr Assoc* 2018;106:142–4.
29 Llullaku SS, Hyseni NS, Bytyçi CI, *et al*. Evaluation of trauma care using TRISS method: the role of adjusted misclassification rate and adjusted w-statistic. *World J Emerg Surg* 2009;4:2.
30 Wang W, Dillon B, Bouamra O. An Analysis of Hospital Trauma Care Performance Evaluation. *J Trauma* 2007;62:1215–22.
31 Younge PA, Coats TJ, Gurney D, *et al*. Interpretation of the Ws statistic: application to an integrated trauma system. *J Trauma* 1997;43:511–5.
32 von Hippel PT. How Many Imputations Do You Need? A Two-stage Calculation Using a Quadratic Rule. *Sociol Methods Res* 2020;49:699–718.
33 Marshall A, Altman DG, Holder RL, *et al*. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57.
34 Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *J R Stat Soc Series B* 1996;58:267–88.
35 Tavakoli N, Saberian P, Bagheri Faradonbeh S, *et al*. Cost-effectiveness Analysis of Air Emergency Versus Ground Emergency Medical Services Regarding the Patient's Transportation and Treatment in Selected Hospital. *Med J Islam Repub Iran* 2022;36:113.
36 Taylor C, Jan S, Curtis K, *et al*. The cost-effectiveness of physician staffed Helicopter Emergency Medical Service (HEMS) transport to a major trauma centre in NSW, Australia. *Injury* 2012;43:1843–9.
37 Hepple DJ, Durrand JW, Bouamra O, *et al*. Impact of a physician-led pre-hospital critical care team on outcomes after major trauma. *Anaesthesia* 2019;74:473–9.
38 Beaumont O, Lecky F, Bouamra O, *et al*. Helicopter and ground emergency medical services transportation to hospital after major trauma in England: a comparative cohort study. *Trauma Surg Acute Care Open* 2020;5:e000508.
39 Elkbuli A, Dowd B, Sanchez C, *et al*. Emergency Medical Service Transport Time and Trauma Outcomes at an Urban Level 1 Trauma Center: Evaluation of Prehospital Emergency Medical Service Response. *Am Surg* 2022;88:1090–6.
40 Littlewood N, Parker A, Hearns S, *et al*. The UK helicopter ambulance tasking study. *Injury* 2010;41:27–9.
41 Chen X, Guyette FX, Peitzman AB, *et al*. Identifying patients with time-sensitive injuries: association of mortality with increasing prehospital time. *J Trauma Acute Care Surg* 2019;86:1015–22.
42 Ghorbani P, Troëng T, Brattström O, *et al*. Validation of the Norwegian survival prediction model in trauma (NORMIT) in Swedish trauma populations. *J Br Surg* 2020;107:381–90.
43 Jones JM, Skaga NO, Søvik S, *et al*. Norwegian survival prediction model in trauma: modelling effects of anatomic injury, acute physiology, age, and co-morbidity. *Acta Anaesthesiol Scand* 2014;58:303–15.
44 Gavrilovski M, El-Zanfaly M, Lyon RM. Isolated traumatic brain injury results in significant pre-hospital derangement of cardiovascular physiology. *Injury* 2018;49:1675–9.
45 Holtenius J, Mosfeldt M, Enocson A, *et al*. Prediction of mortality among severely injured trauma patients. A comparison between TRISS and machine learning-based predictive models. *Injury* 2024;55.
46 Nasser AAH, Khouli Y. The Impact of Prehospital Transport Mode on Mortality of Penetrating Trauma Patients. *Air Med J* 2020;39:502–5.
47 Chen X, Gestring ML, Rosengart MR, *et al*. Speed is not everything: identifying patients who may benefit from helicopter transport despite faster ground transport. *J Trauma Acute Care Surg* 2018;84:549–57.
48 Gavrilovski M, Griggs JE, Ter Avest E, *et al*. The contribution of helicopter emergency medical services in the pre-hospital care of penetrating torso injuries in a semi-rural setting. *Scand J Trauma Resusc Emerg Med* 2021;29:112.
49 Reitala E, Lääperi M, Skrifvars MB, *et al*. Development and internal validation of an algorithm for estimating mortality in patients encountered by physician-staffed helicopter emergency medical services. *Scand J Trauma Resusc Emerg Med* 2024;32:33.
50 O'Neill M, Cheskes S, Drennan I, *et al*. Injury severity bias in missing prehospital vital signs: Prevalence and implications for trauma registries. *Injury* 2025;56:111747.