AWARD NUMBER: W81XWH-19-2-0047

TITLE: Exploitation of Bayesian Networks for Clinical Decision Support on the Battlefield

PRINCIPAL INVESTIGATOR: Col Nigel Tai

CONTRACTING ORGANIZATION: Queen Mary University, London, UK

REPORT DATE: October 2022

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release; Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE
October 2022

2. REPORT TYPE
Annual

3. DATES COVERED
15Sep2021-14Sep2022

4. TITLE AND SUBTITLE

Exploitation of Bayesian Networks for Clinical Decision Support on the Battlefield

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-19-2-0047

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR Colonel Nigel RM Tai

E-Mail: nigel.tai@nhs.net

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Queen Mary University London
327 Mile End Road London
E1 4NS
UK

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

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

U.S. Army Medical Research and Development Command

Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

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

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Purpose: The purpose of this research is to optimize the care of battlefield trauma patients through the development of Bayesian-network (BN) machine learning-powered clinical decision-support (CDS) tools. **Scope:** The scope of the research encompasses the refinement of existing BNs, and development and prototyping of new BNs designed for pre-hospital, en-route, and deployed healthcare facility stages of care, such that CDS prototypes are available for piloting and assessment in future, real-world clinical studies. **Major Findings – Year 3 Data.** Permissions and access to UK Joint Theatre Trauma Registry Transfusion Dataset gained (needed for building transfusion model) and access to US Joint Theatre Trauma System gained (needed for external validation of model set) achieved. UK Human Research Authority permission gained for Decision Support end user use-ability studies. **Model generation.**1. The Trauma Induced Coagulopathy (TIC) BN model was finessed with novel weightings garnered from our previous work on diagnostic accuracy and clinician uncertainty. The updated BN model, which incorporated injury accuracy and clinician uncertainty of injury predictors, was tested with a cohort of civilian patients. TIC-BN Model Performance: Prehospital Discrimination was moderate with Area Under Receiver Operator Curve (AUROC) 0.77 (95% confidence interval (CI) 0.73–0.82). Accuracy was 0.81 and calibration slope was 0.925 with intercept 0.028. Updating the model with Diagnostic Accuracy weightings showed that discrimination was similar with AUROC 0.78 (95% CI 0.74–0.82) and accuracy 0.79. Updating the [Prehospital model weighted by the accuracy] with uncertainty weightings: Discrimination was similar with AUROC 0.78 (95% CI 0.74–0.82) and accuracy 0.78. We concluded that predictive performance was not reduced between iterations and incorporation of measures of uncertainty and diagnostic accuracy may improve clinical validity and user adoption.2 Internal validation of Trauma-induced Acute Kidney Injury (TAKI) model was undertaken using 1234 civilian trauma patients with median age 36, 81% male, median Injury Severity Score 17, 20% penetrating mechanism, mortality 11%. Overall, 32% developed AKI within 3 days of admission, of which 68% were mild and 32% severe. Mortality was 7% and 33% respectively. Internal validation demonstrated excellent performance at ED prediction time point (AUROC 0.93, calibration slope 1.034 and intercept -0.018, accuracy 0.87), as well as excellent performance at the ITU prediction time point (AUROC 0.93, slope 1.020, intercept -0.005, accuracy 0.88). We concluded that an individual patient's risk of TAKI can be reliably predicted from information available at initial assessment as well as following resuscitation. **Software + system engineering.** Prototype Clinical Decision Support System user interface (real time on server, prototype on Figma) including an explanation of the model output in natural language to the user, was tested with end users. R-programming language-based prediction explanation generator environment has been developed and the server technology for integration of BN predictor capture and risk presentation, multiple Apps to interact with the same Server (i.e. improving future scalability). **Conclusion** The project is proceeding toward its endpoint of delivering a suite of validated prototype clinical decision support tools that will enhance the care of injured service men and women.

15. SUBJECT TERMS Trauma, Prediction, Coagulopathy, Mortality, Limb Salvage, Transfusion, Machine Learning, Clinical Decision Support, Military Surgery

16. SECURITY CLASSIFICATION OF"			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRDC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	Unclassified	43	

TABLE OF CONTENTS **Page**

1. INTRODUCTION:

The research subject concerns optimizing the care of battlefield trauma patients through the purposeful development and validation of accurate, clinically credible prognostic tools that will support informed and personalized treatment decisions by quantifying clinically relevant risks. The scope of the research encompasses the development and prototyping of tools designed for pre-hospital, en-route, and deployed healthcare facility stages of care, such that these prototypes are available for piloting and assessment in a clinical context.

2, KEYWORDS:

Trauma, Prediction, Coagulopathy, Mortality, Limb Salvage, Transfusion, Machine Learning, Clinical Decision Support, Military Surgery

3. ACCOMPLISHMENTS:

The scope of this project is to use machine-learning techniques established by the funded research group (COMputer Battlefield Assistance in Trauma care and Injury Decision-support: COMBAT-AID) to enable military trauma clinicians to provide precision-medicine on the battlefield. The date of Project Start was 15 September 2019.

- **Data.** Permissions gained and access supplied to UK MoD Time Stamped Blood Bank Transfusion data concerning 20827 units supplied to 2187 patients within 24hrs of injury. Permission gained for use of US Joint Trauma Theatre System data. Preparing QMUL reception data-vault for latter.
- **Model generation 1**. The Trauma Induced Coagulopathy BN model was finessed with novel weightings garnered from our previous work on diagnostic accuracy and clinician uncertainty. The updated BN model, which incorporated injury accuracy and clinician uncertainty of injury predictors, was tested with a cohort of civilian patients. TIC-BN Model Performance: Prehospital Discrimination was moderate with Area Under the Receiver Operator Curve (AUROC) 0.77 (95% confidence interval (CI) 0.73–0.82). Accuracy was 0.81 and calibration slope was 0.925 with intercept 0.028. Updating the model with Diagnostic Accuracy weightings showed that discrimination was similar with AUROC 0.78 (95% CI 0.74–0.82) and accuracy 0.79. Updating the [Prehospital model weighted by the accuracy] with uncertainty weightings: Discrimination was similar with AUROC 0.78 (95% CI 0.74–0.82) and accuracy 0.78. We concluded that predictive performance was not reduced between iterations and incorporation of measures of uncertainty and diagnostic accuracy may improve clinical validity and user adoption
- **Model development 2**. A key component of the mortality model was built and validated – the Trauma-induced Acute Kidney Injury (TAKI) model. This was undertaken using 1234 civilian trauma patients with median age 36, 81% male, median Injury Severity Score 17, 20% penetrating mechanism, mortality 11%. Overall, 32% developed AKI within 3 days of admission, of which 68% were mild and 32% severe. Mortality was 7% and 33% respectively. Internal validation demonstrated excellent performance at ED prediction time point (AUROC 0.93, calibration slope 1.034 and intercept -0.018, accuracy 0.87), as well as excellent performance at the ITU prediction time point (AUROC 0.93, slope 1.020, intercept -0.005, accuracy 0.88). We concluded that an individual patient's risk of TAKI can be reliably predicted from information available at initial assessment as well as following resuscitation
- **Interface development**. Significant improvements in the design and layout of the prototype Clinical Decision Support (CDSS) user interface were made (real time on server, prototype on Figma) including an explanation of the model output in natural language to the user, was tested with end users. R-programming language-based prediction explanation generator environment has been developed and the server technology for integration of BN predictor capture and risk presentation.
- **End user testing.** Ethics permission (UK equivalent of Institutional Review Board) gained for end-user testing. Pilot study conducted and found that the Clinical Decision Support System achieved a user satisfaction score of 60. Sixty percent of users said the CDSS would improve their clinical performance and would be useful in their clinical practice, and 40% said it would enhance their clinical effectiveness. Potential barriers related to the adopters, organisation and external environment were most prevalent.

What opportunities for training and professional development has the project provided?

Three PhD students are supervised by Col N Tai, Prof K Brohi, Dr William Marsh, Mr Zane Perkins, and Dr Evangalina Kyrimi. All 3 students are registered to Queen Mary University London. Supervisors and students enrol in QMUL programmes to foster their continuing development https://www.qmul.ac.uk/doctoralcollege/?page_id=49. These opportunities are a requirement for satisfactory study and supervision.

How were the results disseminated to communities of interest?

.

Via presentations at national and international academic fora and publication in relevant journals.

What do you plan to do during the next reporting period to accomplish the goals?

1. Complete User Interface Development (transfer from FIGMA design environment to real-time server) informed by End-User Testing; ready interface to include mortality, blood bank and overflight models.
2. Complete mortality model development and external validation on US Joint Theatre Trauma System data.
3. Initiate and complete blood transfusion prediction model.
4. Initiate and complete Overflight model.

4. **IMPACT:**

What was the impact on the development of the principal discipline(s) of the project?

Artificial intelligence and Machine Learning are part of the modern discourse when considering advances in improved medical care, yet many approaches lack the ability to integrate patient data, clinical expertise and evidence-derived knowledge in deriving prognoses that can support decision-making. Published results from our foundational work, contributing to the projects goals, confirm that accurate, Bayesian Network-powered models can be produced that, by incorporating expert knowledge and data-sets, out-perform standard prediction tools with regard to the prediction of traumatic coagulopathy and risk of limb loss in trauma. However, generation of accurate predictions means little if clinicians are reluctant to use the tool, if they do not trust the result or both. We have advanced the discipline by developing a user-configurable interface powered by a 6 level system architecture that marries inputs, computation, outputs and evidence. We have also learnt how the ability of a tool to explain its reasoning in making a prediction is or is not valued by clinicians, and how this influences trust in the model. We have used an novel approach to predicting mortality that incorporates dynamic variables, accounting for the effect of interventions, which provides predictions tailored to the stage of care across a combat area – one that outperforms established mortality prediction tools. And we have made significant in-roads in to the development of a system that predicts likely consumption of blood transfusion resource, based on what the individual patient requires rather than what the clinician is likely to prescribe.

What was the impact on other disciplines?

The research outputs will contribute to the development of clinical decision support tools applicable to non-trauma, non-military patients. The modelling techniques developed are impactful for other areas of human expert decision-making where support is required, as the approaches used are applicable in Law, Engineering, Defence and the Financial sector.

What was the impact on technology transfer?

Nothing to report.

What was the impact on society beyond science and technology?

Nothing to report.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

Nil to report

Actual or anticipated problems or delays and actions or plans to resolve them

Year 3 Staffing issues:

1. Delay in gaining institutional approval to replace Dr Joyner (who has left the team) who had responsibilities for software development. Adverts are out (Oct 2022) to hire a replacement. His duties are now being covered off by Mr Pisirir.

2. Two key members of the research team are either on, or about to take, statutory Maternity Leave. The DoD CDMRP approval of the request for a no-cost additional year extends the project to Sept 2023. One individual returns to the team in April 2023, the other in September 2023. The project extension will assist our ability to deliver project goals according to original timescale. A further no-cost project extension of six months (bringing the end of project date to March 2024) may be required to assure completion. A decision on whether this further no-cost extension will be applied for will be made in January 2023.

Changes that had a significant impact on expenditures

Nil Significant to report. Year 3 spend as below.

~~2021~~	2021 Oct-Dec	2022 Jan-March	2022 April-June	2022 July-Sept	Total Y3
Staff costs	$ 91,566	$ 72,319	$ 71,016	$ 50,628	**$285,529**
Stipend	$ 11,375	$ 5,688	$ 5,688	$ 5,688	**$ 28,439**
PhD fees	$ 0	$ 27,235	$ 0	$ 0	**$ 27,235**
Other non-staff costs	$ 1,445	$ 2,009	$ 502	$ 10,483	**$ 14,439**
Overheads	$ 13,738	$ 13,738	$ 13,738	$ 13,738	**$ 54,952**
Subcontracting costs	$ 0	$ 0	$ 0	$ 0	**$ 0**
TOTAL	$118,124	$ 120,989	$ 90,944	$ 80,536	**$410,593**

Forecast budget for next 2x trimesters.

	2022 Oct - Dec	2023 Jan-March
Staff costs	$62,525	$62,525
Stipend	$5,688	$5,688
PhD fees	$0	$0
Other non-staff costs	$6,500	$5,200
Overheads	$13,738	$13,738
Subcontracting costs	$0	$0
TOTAL	**$88,450**	**$87,150**

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Significant changes in use or care of human subjects

Nil Significant to report.

Significant changes in use or care of vertebrate animals

Nil Significant to report.

Significant changes in use of biohazards and/or select agents

Nil Significant to report

6. PRODUCTS:

- **Publications, conference papers, and presentations**

Journal publications.

Impact of ischaemic duration on lower limb salvage in combat casualties. Perkins ZB, Kersey AJ, White AM et al. Ann Surg 2022; 276: 532-538.

Books or other non-periodical, one-time publications.

Other publications, conference papers and presentations.

Incorporating Clinician Accuracy and Uncertainty of Prehospital Injury Diagnosis Improves Clinical Validity of a Decision Support System Without Compromising Performance of a Bayesian Network to Predict Trauma- Induced Coagulopathy. Jared M Wohlgemut, Evangelia Kyrimi, Andrea Rossetto et al. Poster Presentation. Military Health Systems Research Symposium, Florida. September 2022. Accuracy and Uncertainty of Prehospital Clinical Assessment to Diagnose Major Injuries. Jared M Wohlgemut, Max Marsden, Rebecca Stoner et al. European Society of Emergency Surgery and Trauma Symposium. Oslo. April 2022. Predicting trauma-induced coagulopathy in combat casualties – updating a civilian AI risk model for use in a military population. RS Stoner, E Kyrimi, J Wohlegemut et al. Podium presentation European Society of Emergency Surgery and Trauma Symposium. Oslo. April 2022. Developing an AI prediction model for Trauma-induced Acute Kidney Injury. RS Stoner, E Kyrimi, E Pisirir et al. American Association for the Surgery of Trauma Annual Meeting. Chicago. September 2022.

Website(s) or other Internet site(s)

- **Technologies or techniques**

Nil significant to report

- **Inventions, patent applications, and/or licenses**

Nil significant to report

- **Other Products**

Nil significant to report

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Name	Principle Role	Project time	Contribution	Other funding support	Notes
Nigel Tai	PI	13.3%	Overall leadership functions; review meetings, hiring and staff establishment	MoD employee	Honorary QMUL and NHS contracts
William Marsh	CI	13.3%	Leadership esp wrt Computer Science aspects; review meetings and staff establishment.	QMUL employee	
Zane Perkins	Investigator	20%	Project supervision and academic management of Clinical PhD students; review meetings; hiring and staff establishment	NHS employee	Part funded by DoD grant
Evangelina Kyrimi	Post-Doctoral Research Assistant	100% (currently on parental leave)	Model development and supervision of Computer Science PhD Student; review meetings.	Nil	Fully funded by DoD grant
Rebecca Stoner	QMUL PhD Candidate Clinical Fellow	100%	Model development & refinement	Nil	COVID 19 prevented earlier start
Jared Wohlgemut	QMUL PhD Candidate Clinical Fellow	100%	Clinical Decision Support system Interface development, assessment and workflow integration	Nil	COVID 19 prevented earlier start
Erhan Pisirir	QMUL PhD Candidate	100%			
Javier Sandin Llorente	Operations manager Centre for Trauma Sciences QMUL	20%	Liaison with funder, QMUL Research Office, financial oversight and contractual management	QMUL employee supported by multiple C4TS awards	
Prof Karim Brohi	CI	2.67%	Critical analysis, leadership, authority for Centre 4 trauma science data resources.	Full time QMUL academic	
~~Dr Chris Joyner~~	QMUL Comp Science Lecturer	60% (has left project – funds to be redirected in to replacement hire)	Software and System Developer	Full time QMUL academic	

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nil significant to report

What other organizations were involved as partners?

Nil significant to report

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:

QUAD CHARTS:

9. APPENDICES:

ANNUAL REPORT TECHNICAL NARRATIVE W81XWH-19-2-0047 – YR 3

EXPLOITATION OF BAYESIAN NETWORKS FOR CLINICAL DECISION SUPPORT ON THE BATTLEFIELD

AUTHOR COL NIGEL TAI (PRINCIPLE INVESTIGATOR)

REPORT DATE 13 SEPT 2022 (SUBMITTED 12 OCT 2022)

DETAILED PROGRESS REPORT Year 3.

This report is structured to focus on Year 3 accomplishments as per the Specific Aims and Major tasks template of the original Statement of Work. New data, abstracts and other relevant evidence of progression pertaining to year 3 accomplishments are provided within this report. Materials and evidence pertaining to Year 1 and Year 2 reports are provided in previously submitted annual reports.

Aim/Task 1 Apply for Local IRB and US HRPO study approval – ONGOING 80% COMPLETE.

Year 3 progress. The UK Human Research Authority and Barts Health NHS Trust granted approval (8th June 2022 and 6th July 2022) for human research for the purposes of use-ability studies.

*[“**RE: IRAS 308679. Confirmation of Capacity and Capability at Barts Health NHS Trust.***

***Full Study Title:** Evaluation of Usability, Utility and Credibility of a Clinical Decision Support System for Trauma*
***Site PI:** Mr Nigel Tai*
***Protocol version:** V1.4 17.03.2022*
***Latest HRA Approval date:** 08.06.2022*

*This email confirms that **Barts Health NHS Trust** has the capacity and capability to deliver the above referenced study.*

*Please find attached the signed OID **and** agreed SoECAT as confirmation.*

***Barts Health NHS Trust** agrees to start this study on a date to be agreed when the sponsor gives the green light to begin (as applicable). Please ensure the R&D office (research.governance@qmul.ac.uk) is provided with this date.”]*

Yr 4 Plans: Complete US HRPO approvals to enable use of JTTS data to externally validate models.

Specific Aim 1 – Develop and validate a prognostic model for Trauma-Induced Coagulopathy in injured military personnel (TIC-MIL). COMPLETE

Subtask 1: Literature review and develop structure of TIC-MIL Bayesian Network [Will include a systematic review of causal factors for TIC] COMPLETE

Subtask 2: Construct a development dataset from UK military data sources COMPLETE

Milestone #2 Complete development dataset for the TIC-MIL prognostic model COMPLETE

Subtask 3: Learn parameters of the TIC-MIL Bayesian Network COMPLETE

Subtask 4: Cross-validation and model refinement COMPLETE

Year 3 progress.

1. This work has been progressed as a knowledge product “Methodology for recalibrating TIC civilian model for combat casualties” – accepted by Journal of Biomedical Informatics with revisions, revisions underway, for resubmission December 2022. See Appendix.

2. The TIC model represents one of the major backbone AI engines used for all other modelling outputs in this research. It was built using inputs concerning clinical evaluation of injury and measures of physiology available to the physician working within an Emergency Department setting within a replete civilian or military (Role 3) environment. In order to engineer robustness of prediction in less replete settings (Role 2 or Role 1) it was decided to build in weightings that would allow the model to account for the Accuracy of physician clinical assessment of injuries, and to account for the certainty that the physician has in these assessments, as this would be important for Specific Aims 8 and 9 (which address conversion of the models to desired clinical decision support systems) as well as Specific Aim 1 (model refinement). The work is presented in this section of the report to facilitate cross-Aim continuity.

This work was begun in Year 2 and completed in Year 3. This structural improvement (Refinement of Model) was informed by a retrospective study evaluating the diagnostic accuracy of clinical examination to detect major injuries and bleeding, and the effect of clinician uncertainty on diagnostic accuracy. The abstract of this work is below, which was presented at ECTES in April 2022 (European Congress of Trauma and Emergency Surgery, Oslo, Norway). The manuscript of this work will be submitted in October 2022, and represents novelty in that very few predictive decision support tools account for clinical inaccuracy and uncertainty. Two abstracts cover this work – one representing the base-level research in to accuracy and uncertainty, and one covering the exploitation of this knowledge to refine the model. It was presented at MHSRS September 2022 (Military Health Services Research Symposium, Kissimmee, Florida, USA), and won an “Honorable Mention” at the Poster Prize session.

Abstract: Timely and accurate identification of life- or limb-threatening injuries (LLTIs) is a fundamental objective of trauma care that directly informs triage and treatment decisions. However, the diagnostic accuracy of clinical examination to detect LLTIs is largely unknown.

Objective: To assess the diagnostic accuracy of initial clinical examination for detecting LLTIs in trauma patients pre-hospital, and to identify factors that affect accuracy.

Design: Diagnostic accuracy study conducted among a retrospective cohort of 947 trauma patients between January 2019 and December 2020.

Setting: Major Trauma Centre (MTC) within the London Trauma System.

Participants: Consecutive adult (≥16 years old) trauma patients who were examined at the scene of injury by an experienced trauma clinician and admitted to one MTC were included. Patients were excluded if they were <16 years old, uninjured, or suffered a thermal injury.

Main Outcomes and Measures: Sensitivity, False Negative Rate (missed injury), False Positive Rate (over-diagnosis) and Positive Predictive Value (PPV) were calculated for each LLTI. Diagnostic imaging, operative and post-mortem findings were used as the reference standard.

Results: Among 947 trauma patients (821 male [86.7%]; median age 31 years [range 16-89]; 569 blunt mechanisms [60.1%]), 522 (55.1%) had a LLTI. Clinical examination had a moderate ability to detect LLTIs: life-threatening head injury (sensitivity 69.7%, PPV 59.1%), life-threatening chest injury (sensitivity 58.7%, PPV 53.3%), life-threatening abdomen injury (sensitivity 51.9%, PPV 30.7%), and life-threatening pelvic injury (sensitivity 23.5%, PPV 50.0%). Clinical examination had a poor ability to detect life-threatening thoracic (sensitivity 31.9%, PPV 29.0%) and abdominal (sensitivity 42.4%, 23.3%) bleeding. Missed injuries were more common in patients with multiple injuries (odds ratio (OR) 1.83, 95% confidence interval (CI) 1.62-2.07) or shock (hypotension OR 0.993, 95% CI 0.988-0.998; tachycardia OR 1.01, 95% CI 1.00-1.01). Over-diagnosis was

more common when there was diagnostic uncertainty (OR 6.42, 95% CI 4.63-8.99) or hypotension (OR 0.991, 95% CI 0.986-0.995).

Conclusions and Relevance: Clinical examination, even when performed by experienced trauma clinicians, had a moderate ability to detect LLTIs. Polytrauma, shock, and diagnostic uncertainty worsened accuracy. The accuracy of clinical examination findings should be considered when making clinical decisions.

Figures and Tables:

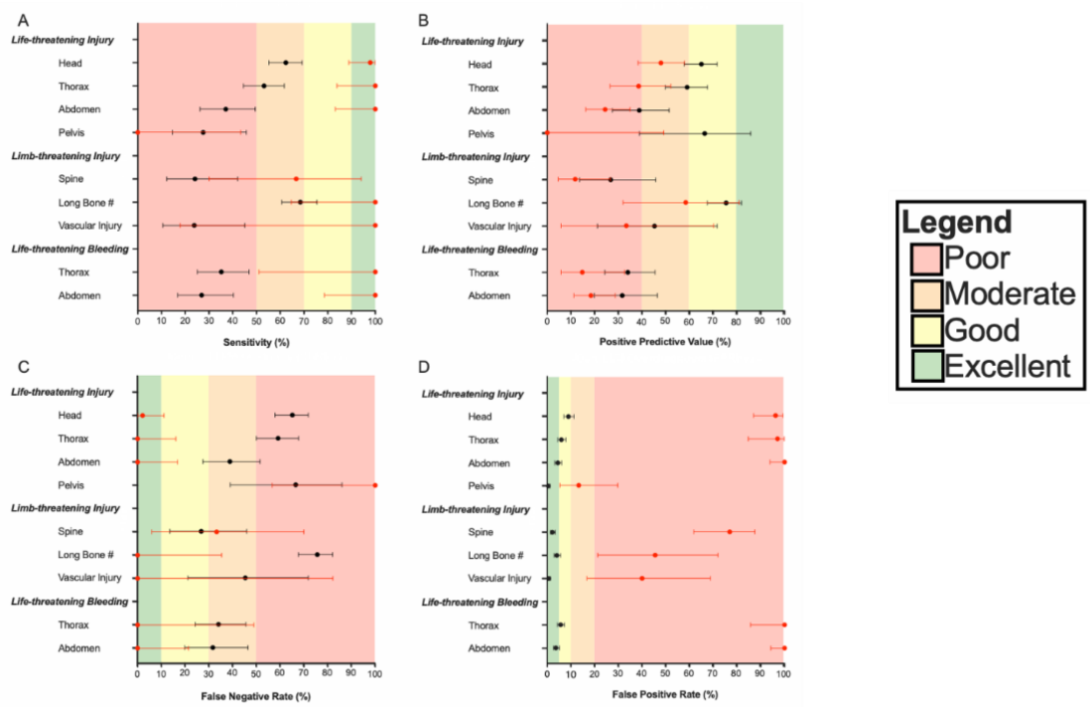


Fig 1. Measures of diagnostic accuracy of clinical examination to diagnose life- and limb-threatening injuries, including a) sensitivity, b) positive predictive value, c) false negative rate, d) false positive rate. Black dots are measures of accuracy of injury diagnoses when clinicians were certain (with 95% confidence intervals). Red dots were when clinicians were uncertain. The vertical colours represent a consensus of acceptable levels of diagnostic accuracy. **Explanation:** clinicians have only a moderate to poor ability to detect major injuries of the head, torso, and major bleeding. When clinicians are uncertain, their sensitivity increases, while their positive predictive value decreases, and their false positive rate increases.

	Missed Injuries (n=265)				Overdiagnosed Injuries (n=380)			
	Univariate		Multivariate		Univariate		Multivariate	
Parameter	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p
Patient factors								
Age	1.00 (0.991 to 1.01)	0.97			0.998 (0.990 to 1.01)	0.59		
Female Sex	1.46 (0.973 to 2.16)	0.06	1.19 (0.743 to 1.87)	0.47	0.744 (0.498 to 1.10)	0.14		
Penetrating MOI	0.704 (0.522 to 0.944)	0.02	1.06 (0.724 to 1.54)	0.78	0.777 (0.594 to 1.01)	0.06	0.940 (0.675 to 1.31)	0.72
Polytrauma	1.86 (1.66 to 2.08)	<0.001	1.83 (1.62 to 2.07)	<0.001	1.18 (1.07 to 1.30)	<0.001	1.12 (0.999 to 1.25)	0.05
PH GCS	0.961 (0.929 to 0.994)	0.02	0.999 (0.957 to 1.04)	0.97	0.991 (0.960 to 1.02)	0.57		
PH SBP	0.992 (0.988 to 0.997)	<0.001	0.993 (0.988 to 0.998)	0.005	0.991 (0.987 to 0.995)	<0.001	0.991 (0.986 to 0.995)	<0.001
PH HR	1.01 (1.00 to 1.01)	0.006	1.01 (1.00 to 1.01)	0.03	1.0 (0.995 to 1.00)	0.88		
Clinician factors								
Spec: Anaesthetics	1.14 (0.838 to 1.55)	0.4			0.628 (0.470 to 0.836)	.002	0.873 (0.627 to 1.21)	0.42
Intensive Care	1.04 (0.603 to 1.74)	0.89			0.568 (0.339 to 0.930)	0.03	0.707 (0.386 to 1.26)	0.25
Diagnostic Uncertainty	0.921 (0.670 to 1.26)	0.621			6.65 (4.89 to 9.12)	<0.001	6.42 (4.63 to 8.99	<0.001
Environment factors								
Nightshift	0.959 (0.722 to 1.28)	0.77			0.74 (0.570 to 0.960)	0.02	0.814 (0.600 to 1.10	0.18

Table 3 footnote: Model statistics represented are Odds Ratios (95% confidence intervals), p values.
Referents in the model: for Female Sex was Male, for Penetrating MOI was Blunt, for Anaesthetics and ICU base specialty was ED, for Clinician Diagnostic Uncertainty was Certain, and for Nightshift was Dayshift.
MOI= mechanism of injury; GCS= Glasgow coma scale; SBP= systolic blood pressure; HR= heart rate; Spec= base specialty; Polytrauma = continuous variable of number of abbreviated injury scale (AIS) categories injured.

Table 1. Univariate and multivariate logistic regression analyses, with missed life- and limb-threatening injuries (false negative, n=265), and overdiagnosis (false positives, n=380) as dependent variables, respectively. **Explanation** Factors which significantly increase the likelihood of missing injuries include if the patient is multiply injured, or shocked. Factors which significantly increase the likelihood of overdiagnosis is uncertainty and shock.

Abstract. Traumatic hemorrhage is the leading preventable cause of death after injury. Trauma-induced coagulopathy (TIC) is a complication of severe traumatic hemorrhage which is associated with increased mortality. Early treatment of TIC can be life-saving, but identification of at-risk patients can be difficult, especially in prehospital and austere settings. The likelihood of TIC calculated by a clinical decision support system (CDSS) could aid decisions about treatment, transport, and prioritization. The use of such CDSS may result in improved patient outcomes from more timely and informed interventions. A Bayesian Network (BN) has been developed to predict trauma-induced coagulopathy (TIC-BN). BNs are graphical probabilistic models which can be designed using expert knowledge and clinical data. The TIC-BN model incorporates patient demographic, physiological, laboratory, and injury variables. However, the accuracy of prehospital clinical assessment to diagnose injuries may be imperfect, and clinician uncertainty may further reduce diagnostic accuracy. We hypothesized that incorporating clinician accuracy and uncertainty of injury diagnoses in the TIC-BN structure would improve clinical validity by ensuring the CDSS can handle real-world information from users in forward and en route combat casualty care environments, without reducing predictive accuracy.
Our aims were to 1) determine the diagnostic accuracy of prehospital clinical assessment by expert trauma clinicians in identifying major injuries; 2) determine the frequency of clinician uncertainty of major injuries, and the impact on diagnostic accuracy; and 3) determine how incorporating diagnostic accuracy and uncertainty in an updated TIC-BN model structure impacts predictive performance.

Materials and Methods Study Design: Retrospective validation study incorporating clinician diagnostic accuracy and uncertainty nodes into the structure of an existing TIC-BN model. The study has been reported according to STARD and TRIPOD guidelines. Study Population: Consecutively injured patients assessed by expert prehospital trauma clinicians in an urban helicopter emergency medical service and admitted to the regional Level 1 Trauma Center between 01/01/2019 and 12/31/2020 were included. Patients were excluded if the model outcome was unknown (TIC defined as international normalized ratio (INR) >1.2), age <16 years, and non-trauma or thermal injury.Data Sources: Patient demographics, physiology, injuries, and clinician uncertainty of injury diagnoses were collected from prehospital documentation. Laboratory data and discharge diagnoses were collected from the Trauma Center clinical registry. Classification of true outcome (TIC) was facilitated by comparing an expectation-maximization algorithm to the INR results. Discrepancies were resolved by expert consensus. Injury Accuracy and Uncertainty: For each injury, prehospital diagnosis was compared to discharge diagnosis. 2x2 contingency tables were constructed, and sensitivity, and specificity were reported. Contingency tables were repeated for each injury for which clinicians documented diagnostic uncertainty. Updating the TIC-BN structure: BN is a directed acyclic graph with qualitative and quantitative parts. The qualitative BN structure comprises nodes representing random variables and directed arcs signifying causal or influential relationships. The qualitative BN parameters comprise a set of conditional probability functions. Variables of TIC-BN relate to injury (energy, mechanism, Glasgow Coma Scale, hemothorax, abdominal bleeding, unstable pelvis fracture, and long bone fracture), perfusion (lactate, base excess, pH, prehospital heart rate, systolic blood pressure, and temperature), and coagulopathy (prehospital fluid administration). The previously developed TIC-BN had excellent predictive performance at external validation: area under the receiver operator curve (AUC) 0.95, calibration slope 0.96 (ideally 1), Brier Score (BS) 0.05 (ideally 0), and Brier Skill Score (BSS) 0.46 (ideally 1). The TIC-BN structure was revised to incorporate prehospital diagnostic accuracy and clinician uncertainty for hemothorax, abdominal bleeding, and unstable pelvis; and diagnostic accuracy only for long bone fracture, as uncertainty was infrequent. TIC-BN Model Performance: The predictive performance of the updated TIC-BN model was assessed by its discrimination, accuracy and calibration, in three iterations. The ‘prehospital’ iteration included prehospital clinical assessment in injury nodes that represent in hospital diagnostic tests. The ‘prehospital + accuracy’ iteration added nodes for prehospital clinical assessment of injuries. The ‘prehospital + accuracy + uncertainty’ iteration added nodes for clinician uncertainty. Statistical Analysis: Diagnostic accuracy analysis was conducted using Prism v9.0.2. TIC-BN model reasoning was conducted using AgenaRisk and the model’s performance was evaluated using R v4.1.2.

Results Study Population: Our validation cohort comprised 811 patients who met eligibility criteria. Their median age was 32 years (interquartile range (IQR) 23-48), 703 (86.7%) were male, 497 (61.3%) suffered blunt mechanism, median injury severity score (ISS) was 13 (IQR 5-25), and mortality was 69/811 (8.5%). The rate of TIC in our cohort was 82/811 (10. 1%).Injury Accuracy and Uncertainty: The sensitivity and specificity of prehospital clinical assessment were respectively 32.4% and 92.5% for hemothorax; 42.2% and 89.2% for abdominal bleeding; 19.2% and 99.0% for unstable pelvis; and 65.5% and 95.5% for long bone fracture. Clinicians expressed uncertainty in 21/79 of hemothorax (26.6%), in 66/108 of abdominal bleeding (61.1%), in 3/13 of unstable pelvis (23.1%), and in 8/109 of long bone fracture diagnoses (7.3%). When clinicians were uncertain, the sensitivity and specificity of prehospital clinical assessment were respectively 100.0% and 0.0% for hemothorax; 93.3% and 3.7% for abdominal bleeding; 0% and 89.3% for unstable pelvis; and 100.0% and 33.3% for long bone fracture. TIC-BN Model Performance: Prehospital Discrimination was moderate with AUC 0.77 (95% confidence interval (CI) 0.73–0.82). Accuracy was 0.81, the BS was 0.08, and the BSS was 0.14. Calibration slope was 0.925, and calibration intercept was 0.028. Prehospital + Accuracy: Discrimination was similar with AUC 0.78 (95% CI 0.74–0.82), accuracy 0.79, BS 0.08, BSS 0.16, calibration slope 0.922, and calibration intercept 0. 018.Prehospital + Accuracy + Uncertainty: Discrimination was similar with AUC 0.78 (95% CI 0.74–0.82), accuracy 0.78, BS 0.08, BSS 0.15, calibration slope 0.905, and calibration intercept 0.018. Conclusions: Prehospital clinical assessment by expert trauma clinicians to diagnose major injuries is imperfect. Clinician uncertainty worsens accuracy by reducing specificity, due to more false positive diagnoses. In the prehospital context - compared to the in-hospital context in which the original model which was developed and validated - discrimination reduced, though accuracy and calibration remained excellent. By incorporating prehospital accuracy and uncertainty to the TIC-BN structure, predictive performance was not reduced between iterations. The rationale for incorporating this information in the CDSS was to allow users to express their uncertainty, mirroring the cognitive reality of prehospital and austere clinical practice, and take into account known diagnostic inaccuracies, improving its clinical validity.

Figures and Tables

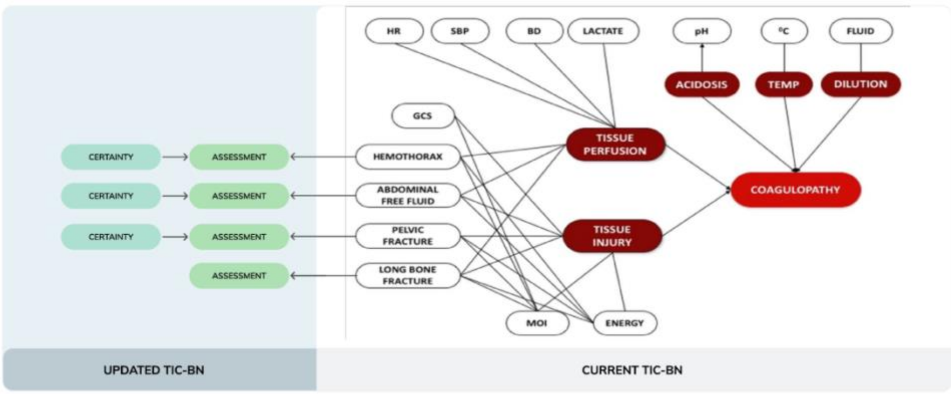


Fig 2- Original BN (on right), published by Perkins et al (Ann Surg 2020), with additional nodes added for accuracy of clinical assessment of injuries, and clinician certainty (on left). In order to test how the addition of nodes for clinician uncertainty and diagnostic accuracy affect BN model performance, these nodes were added to an existing model.

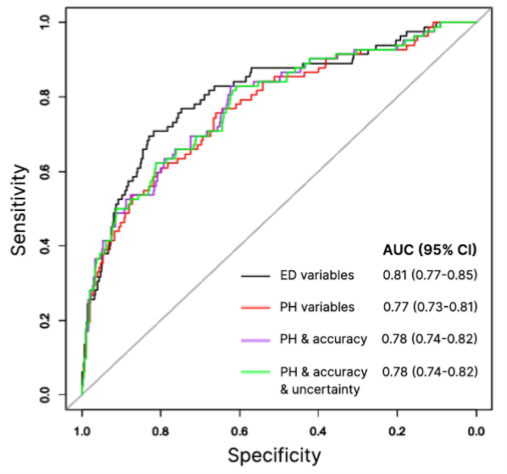


Fig 3. Discrimination shown using area under receiver operator curve, of updated BN model, with iterations relating to data (ED or pre-hospital) and nodes used (diagnostic accuracy, and clinician uncertainty). There is no significant difference in the discriminatory ability of the model, whether it utilises data from the emergency department, pre-hospital, or incorporated clinician uncertainty or diagnostic accuracy.

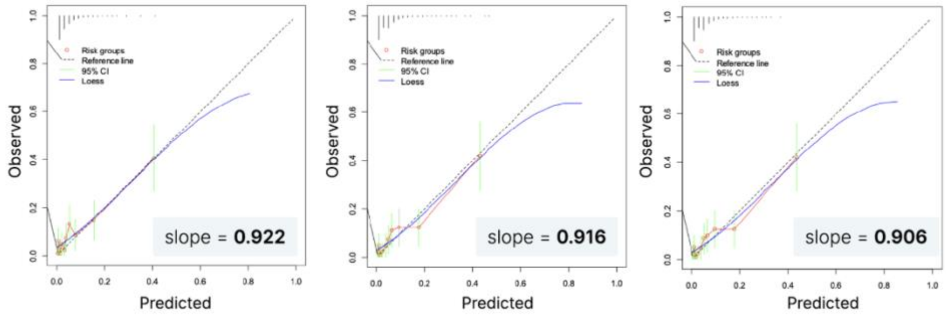


Fig 4. Calibration curves of each iteration. a) PH variables. b) PH & accuracy. c) PH & accuracy & uncertainty. Each iteration demonstrates an acceptable level of calibration, when plotting observed to predicted outcomes.

This study demonstrated that incorporating uncertainty and diagnostic accuracy did not alter the predictive performance of the BN model, in terms of discrimination, calibration and accuracy. The next steps, therefore, are to determine whether there is any compelling indication to either include or exclude incorporating uncertainty and diagnostic accuracy into the model input. On one hand, including these variables may improve user trust in the model as it matches better with their cognitive reality of decision-making in this context, though it will add complexity to the model and time to complete user input. On the other hand, excluding these variables will simplify the model, but may leave the user in a situation where they may calculate the output multiple times if they are uncertain about certain information.

Major Task 2 – External validation of TIC-MIL prognostic performance – ONGOING 20% COMPLETE

Year 3 progress. External validation of TIC-MIL is planned, utilising US DODTR data. A data sharing agreement has been completed for the acquisition of this data. Arrangements are being finalised for the set-up of a secure virtual machine environment in which the data can be hosted by Queen Mary University of London's Data Safe Haven (co-ordinated by the Barts Cancer Centre). This Safe Haven is an infrastructure that has a proven level of data security and suitable information governance processes in place. This is subject to annual evaluation of compliance, which is a complex process involving the auditing of people, processes and technology. This is managed by NHS Digital, through their Information Governance Toolkit programme. This Safe Haven environment will enable us to receive (via secure FTP transfer), analyse, and utilise the data for modelling validation and testing without it ever leaving the secure space. The modelling validation and testing results can be exported for publication and sharing. The set-up of the environment and receipt of the data is anticipated Sept-Dec 2022.

Subtask 1: Construct external validation dataset from US military data sources NOT INITIATED

Milestone #3 Complete validation dataset for the TIC-MIL prognostic model

Subtask 2: External validation of TIC-MIL prognostic performance NOT INITIATED

Milestone #4 Validated prognostic model for Trauma-Induced Coagulopathy

Year 3 update. Contingent on US JTTR Dataset. Plan to validate model in Year 4.

Specific Aim 2 – Develop and validate adaptation of LIMB-MIL prognostic model that can accurately predict projected viability of injured limb, and outcome of limb reperfusion, at initial wound evaluation. COMPLETE

Year 3 progress. COMPLETED, as per previous reports. Outputs from *Specific Aim 2, Major Task 3, Subtask 3: Learn parameters of the new LIMB-MIL Bayesian Network* were presented and published in 2022 as **Impact of ischemia duration on lower limb salvage in combat casualties** was presented at the American Surgical Association (ASA) 142nd annual meeting on 7 April 2022 in Chicago, US and published as Perkins ZB, Kersey AJ, White JM, Lauria AL, Propper BW, Tai NRM, Rasmussen TE. **Impact of Ischemia Duration on Lower Limb Salvage in Combat Casualties**. Ann Surg. 2022 Sep 1;276(3):532-538. doi: 10.1097/SLA.0000000000005560 (See appendix). The dataset developed in this work was exploited to further refine the understanding of time to revascularization for injured limbs as a factor for success or failure in limb salvage. We identified the duration of limb ischaemia as one of the most influential variables on the outcome of a severe lower limb injury. This study describes the relationship between ischaemia duration and limb salvage in combat casualties. This information was incorporated in the LIMB-MIL Bayesian Network.

Introduction: The 6-hour threshold to revascularization of an ischemic limb is ubiquitous in the trauma literature, however, contemporary evidence suggests that this threshold should be less. This study aims to characterize the relationship between the duration of limb ischemia and successful limb salvage following lower extremity arterial trauma.

Methods: This is a cohort study of the United States and UK military service members injured while serving in Iraq or Afghanistan between 2003 and 2013. Consecutive patients who sustained iliac, femoral, or popliteal artery injuries, and underwent surgery to attempt revascularization, were included. The association between limb outcome and the duration of limb ischemia was assessed using the Kaplan-Meier method. **Results:** One hundred twenty-two patients (129 limbs) who sustained iliac (2.3%), femoral (56.6%), and popliteal (41.1%) arterial injuries were included. Overall, 87 limbs (67.4%) were successfully salvaged. The probability of limb salvage was 86.0% when ischemia was ≤1 hour; 68.3% when between 1 and 3 hours; 56.3% when between 3 and 6 hours; and 6.7% when >6 hours ($P<0.0001$). Shock more than doubled the risk of failed limb salvage [hazard ratio=2.42 (95% confidence interval: 1.27-4.62)].

Conclusions: Limb salvage is critically dependent on the duration of ischemia with a 10% reduction in the probability of successful limb salvage for every hour delay to revascularization. The presence of shock significantly worsens this relationship. Military trauma systems should prioritize rapid hemorrhage control and early limb revascularization within 1 hour of injury

Figures and Tables

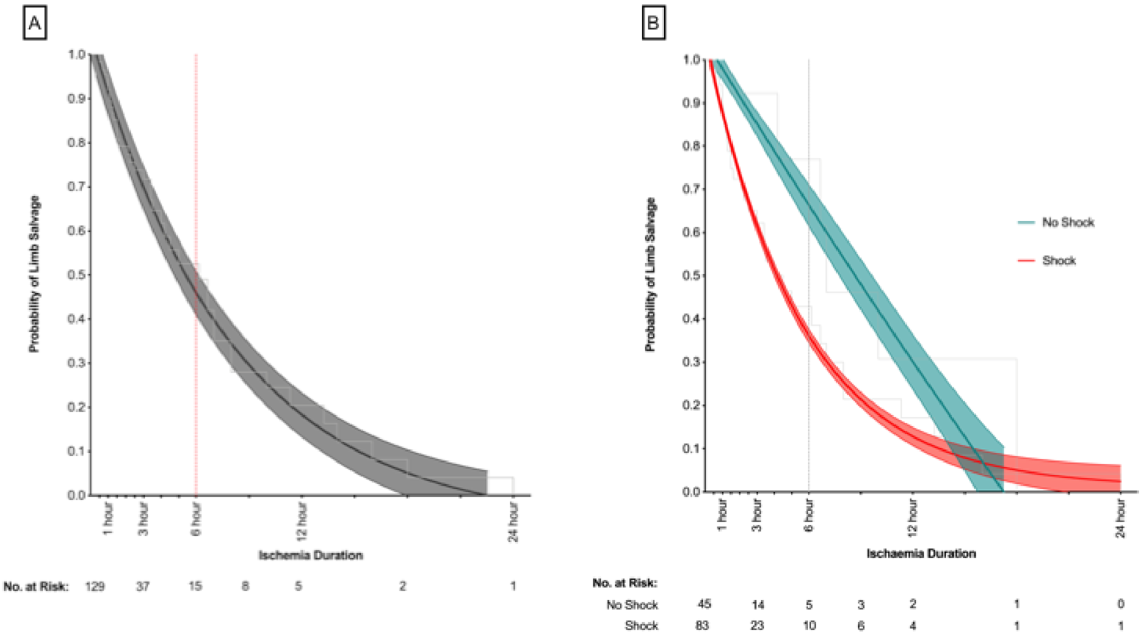


Figure 5 A and B: The relationship between the duration of limb ischemia and outcome. A) The Kaplan-Meier plot estimates the probability of limb salvage with increasing duration of ischaemia in 129 injured limbs that underwent surgical revascularisation. A curve with 95% CIs is fitted to the Kaplan-Meier plot using the least squares method. B) The relationship between the duration of limb ischemia and outcome in casualties presenting with shock (red) or normal haemodynamics (green). The Kaplan-Meier plot estimates the probability of limb salvage with increasing duration of ischemia in 45 injured limbs of casualties with normal haemodynamics and 83 injured limbs of casualties in shock that underwent surgical revascularization. Curves with 95% CIs are fitted to the Kaplan-Meier plots using the least squares method.

Specific Aim 3 - Develop a prototype CDS tool for use of the TIC-MIL and LIMB-MIL prognostic models. ONGOING 65% COMPLETE

Subtask 1: Determine most effective methods of presenting raw probability outputs to support precision medicine. [User simulation experiments and interviews] ONGOING 55% COMPLETE

Subtask 2: Develop user-interface to capture information and present probabilities. ONGOING 70% COMPLETE

Subtask 3: Develop evidence browser and explanation generator for TIC-MIL ONGOING 75% COMPLETE

Subtask 4: Develop evidence browser and explanation generator for LIMB-MIL ONGOING 50% COMPLETE

Milestone #8 Functional evidence browsers for TIC-MIL and LIMB-MIL

Milestone #9 Prototype tool for clinical use of TIC-MIL and LIMB-MIL

Specific Aim 3: Prototype CDSS development (EP, JW)

Year 3 progress: A Bayesian decision support system for our clinical prediction models is developed and hosted on the Queen Mary University of London servers (http://newton.eecs.qmul.ac.uk/). This system, called Bayesian Network Decision Support (BNDS), is a complete overhaul of previous trauma decision support system we had developed and used. BNDS will be used to host and run COMBAT-AID models for trauma decision support. It is developed using Node.js with Express, R, and MySQL for the back-end of the server and Vue.js for the front-end and the user webpage. The developed back-end server is the central part of the system – it handles requests (in JSON format) from users of the system (client), processes those requests and performs actions such as making calls to the R server for model inferencing or collecting data from the database. Every request is followed by a response back to the client, again in JSON format (there are no HTML request/responses). High level overview of the back-end server is shown in Figure 6:

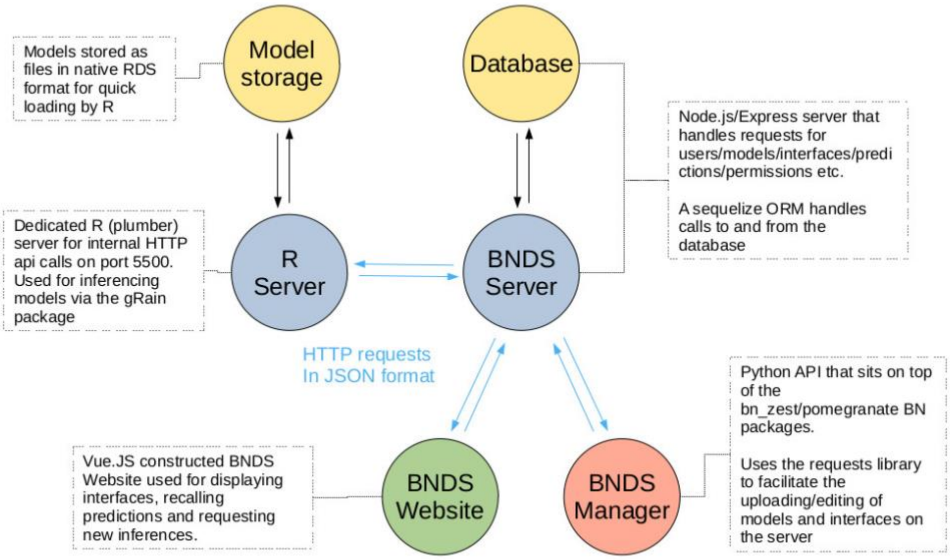
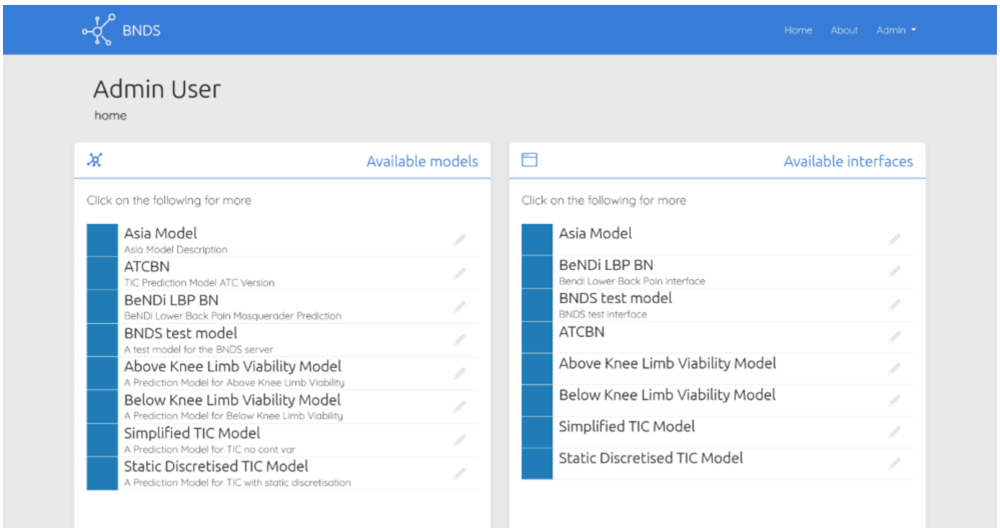
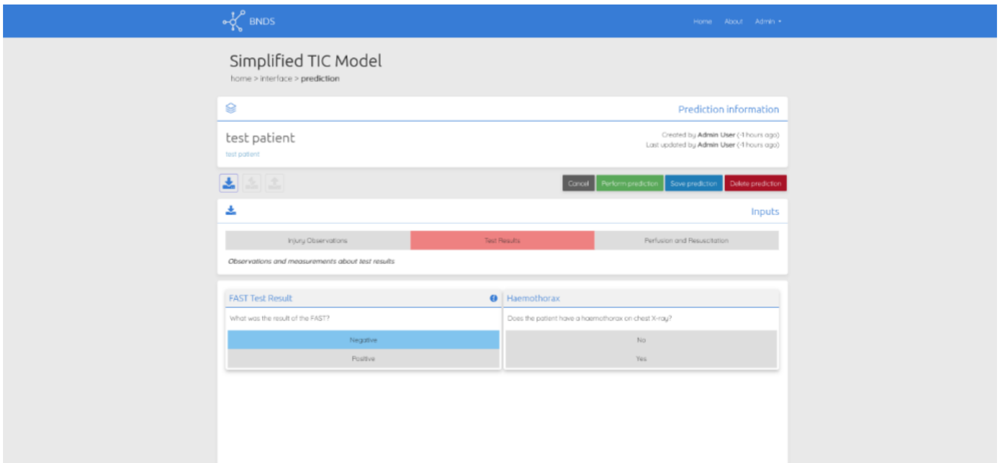


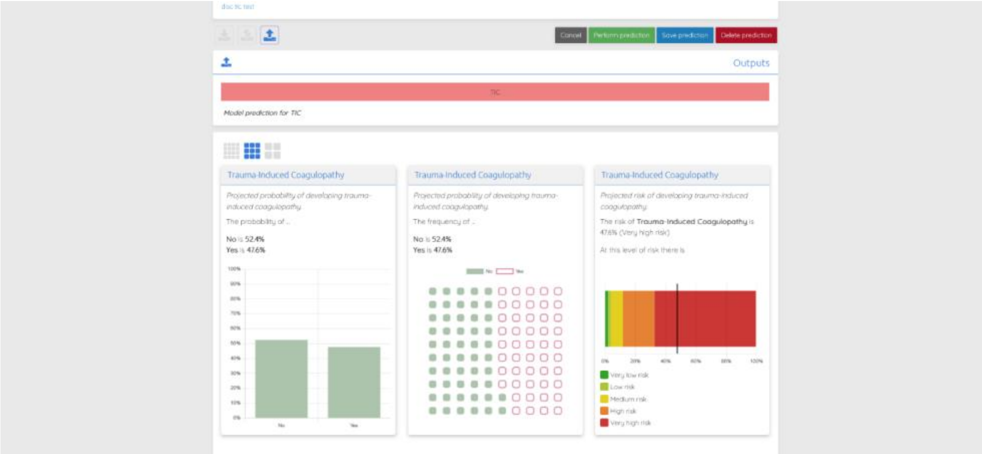
Fig 7. The front-end of the BNDS is the webpage users will engage in order to run new predictions on the prediction models. Currently the main user page looks like the following Figure and allows user to pick a model and enter a new case for prediction.



The input page currently looks like the following Figure (8).



The output page currently looks like the following figure (9). Currently output page displays three visualisation options (bar chart, icon array, and risk stratification) on the same page.



The BNDS server is live and running, in addition to the server we developed a new prototype interface which is used in the usability evaluation of our developed tool as well as an R-based environment to create explanations for the predictions the system provides. The interface prototype and the explanation environment are going to be integrated into the live server once the testing and evaluation process is completed. The former will impact the front-end of BNDS whereas the latter will improve the capabilities of the back-end.

The interface prototype is developed in the prototyping environment Figma, and is disconnected from the back-end server at the moment. The prototype development is also strongly related to *Specific Aim 4: Clinical usability evaluation of developed CDSS* and the development is done incrementally. The improvements from the first prototype and the current one are described under *Specific Aim 4*. Here the current prototype for the interface is shown. Once the evaluation study mentioned under *Specific Aim 4* is completed, the prototype will move to the coded front-end of live BNDS server.

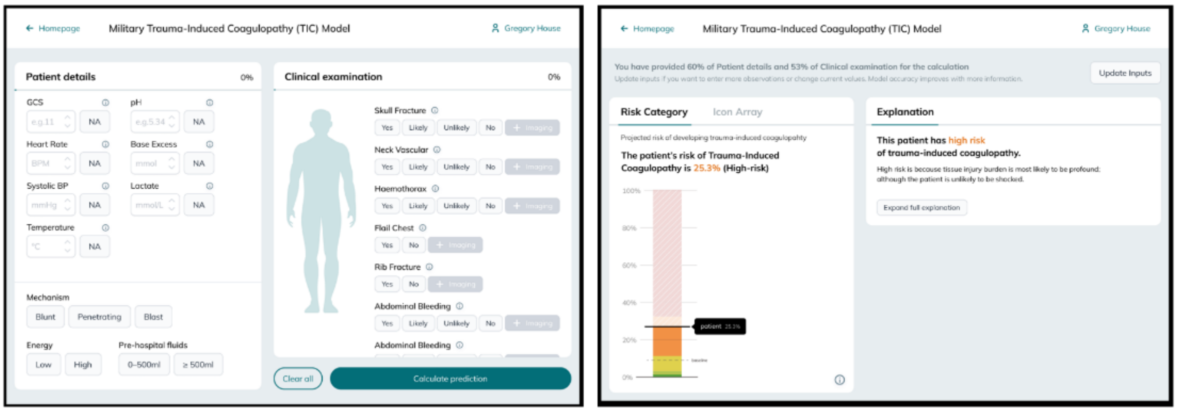
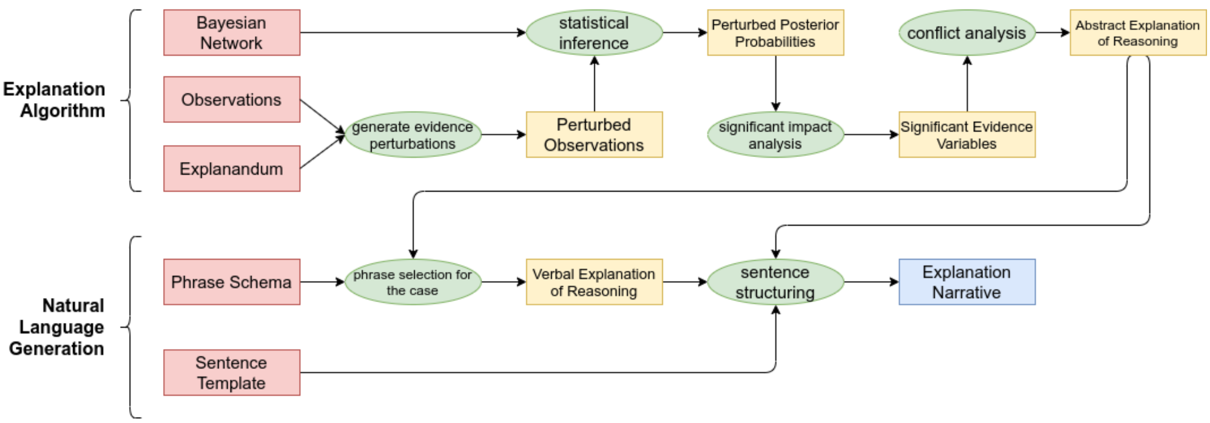
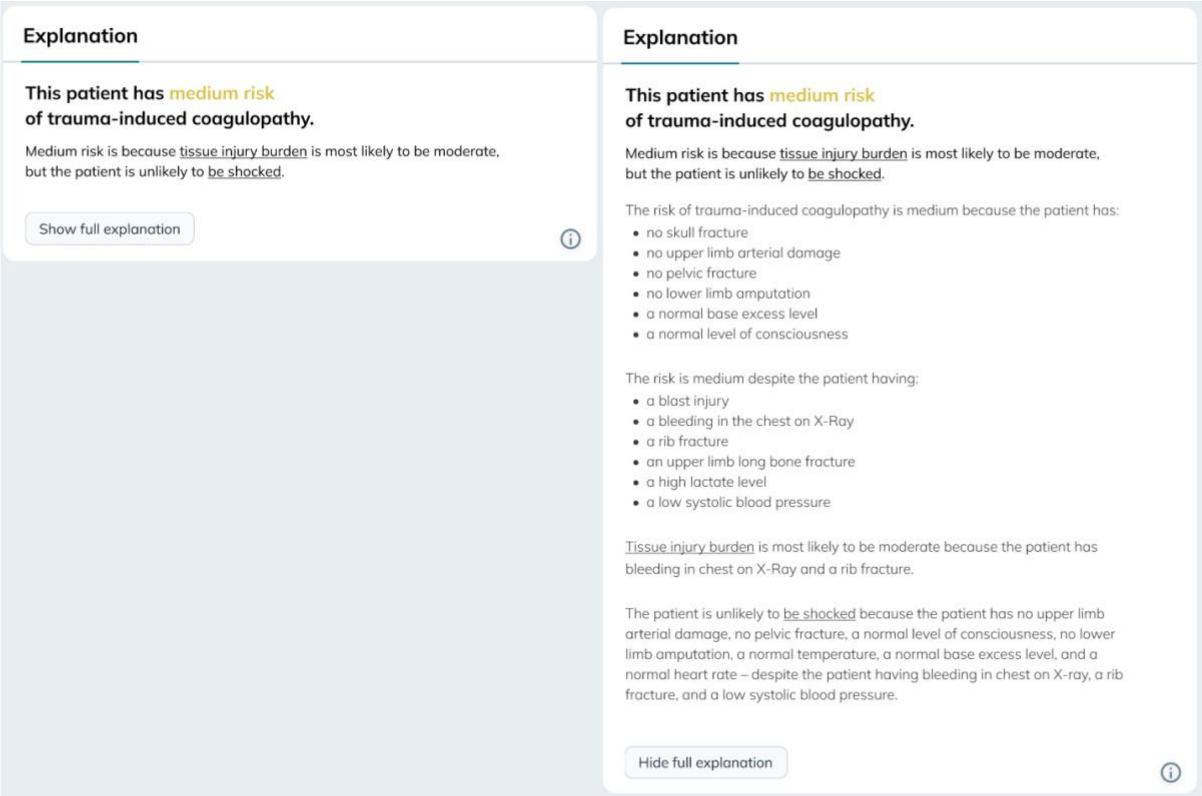


Fig 10. Iteration of user interface using FIGMA

The prediction explanation environment is developed in R, and it uses an extended model structure that we call "explainable Bayesian networks" to generate prediction explanations in natural language. The process flowchart of the explanation environment, including the inputs it takes and the sub-tasks of our incremental explanation generation and narrative creation algorithm, is shown in Figure 11:

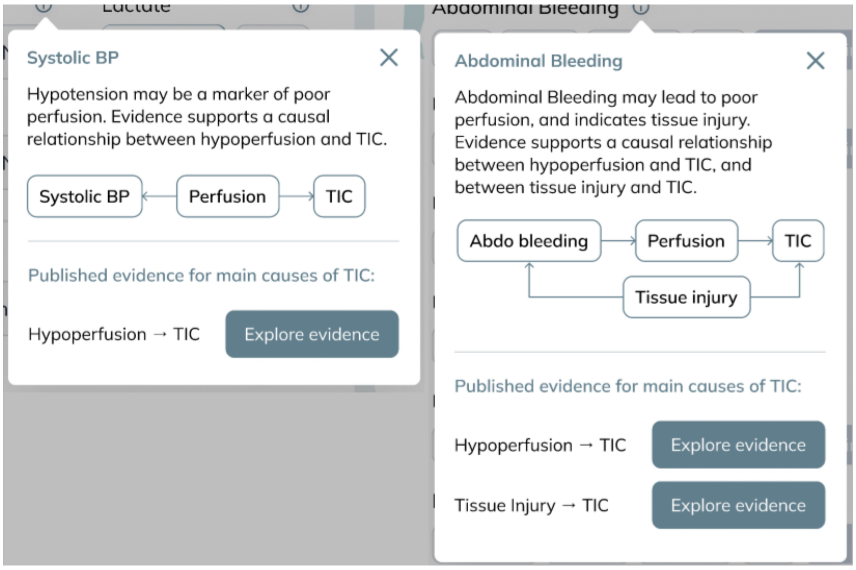


An explanation generated by our environment for a TIC prediction, shown in two parts (short explanation using the latent causes of TIC and the full explanation using all provided observations), is given in the following Figure (12). This is how the explanation of the prediction is displayed in our interface prototype.



The source code for the prediction explanation environment is stored on QMUL enterprise research GitHub page (private repository: https://github.research.its.qmul.ac.uk/acw552/BayesianExplanation).

Once the development and testing is complete, the explanation capabilities will be integrated into the live BNDS server and the prediction output page will include the short and the expandable full explanations. With the prediction explanation generation, the aim is increased trust and transparency in the system. Another branch of work to achieve this aim is explaining the underlying model structure and the medical evidence that supports it. Previously standalone "evidence browser" pages were used, these pages needed improvements. The current design of the improved model evidence is planned to be integrated into the rest of the BNDS system. The integrated evidence browser is a part of our new interface prototype. The design moves away from showing the entire model structure to the end user in favour of selecting the relevant and important parts to link the prediction model with the medical evidence. On the input page the model variables are explained as the following Figure (13).



The integrated evidence browser is based on statements about parts of the model and links to the medical evidence behind the main structure of the model. For example, if user chooses to “explore evidence” behind the relationship between tissue injury and TIC, the system leads the user to the evidence base (Figure 14):

Evidence Browser : Tissue Injury

Tissue Injury → TIC

Published evidence supporting a causal relationship between Tissue Injury and TIC.

First Author, Year	Design	Study Population	Inclusion Criteria	Sample Size	Coagulation Measurement	Risk of Bias (9=low risk)
HUMAN OBSERVATIONAL STUDIES						
Brohi, 2003	RC	SC, Civilian, all MOI	Admitted by Helicopter emergency service	1867	PT, aPTT. TT	6
Brown, 2012	PC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1877	INR	7
Cap, 2011	RC	MC, Military, all MOI	iTBI	1609	INR	8
Carrick, 2005	RC	SC, Civilian, Blunt Trauma	iTBI, GCS <14	184	PT, PPT, Plats	7
Cheddie, 2013	RC	SC, Civilian, all MOI	BD >-2	28	INR	6
Chhabra, 2013	PR	SC, Civilian, all MOI	iTBI, GCS <13, admitted neurosurgery	208	PT, aPPT	6
Cohen, 2009	PC	SC, Civilian, all MOI	TA	168	INR	6
Cohen, 2010	PC	SC, Civilian, all MOI	TA	168	INR	6
Cohen, 2012	PC	SC, Civilian, all MOI	Highest TA	203	INR	8
Cohen, 2013	PC	MC, Civilian, all MOI	Blood transfusion within 6hrs	1198	INR, PTT	9
Cosgriff, 1997	PC	SC, Civilian, all MOI	>10u PRBC/24hr	58	PT, PTT	8
Davis, 1996	RC	SC, Civilian, all MOI	ABG within 1hr of arrival	2954	PT, PTT, Plats, Clinical	7
Deras, 2014	RC	SC, Civilian, all MOI	Admitted to ICU	663	PT, aPPT	6
Dunbar, 2009	CC	SC, Civilian, all MOI	Blood sample within 1hr of ED arrival	42	INR, PT	4

As all the testing and evaluation is completed, the live BNDS server, the new interface prototype, the prediction explanation generator and the model evidence browser will be integrated as the trauma decision support system.

Specific Aim 4 – Evaluate the clinical usability of the prototype CDS tool ONGOING 65% COMPLETE

Major Task 1 – User simulation experiments and user interviews

Subtask 1: Assess clinical usability of the user-interface for predictor entry ONGOING 70% COMPLETE

Subtask 2: Assess clinical usability of user-interface at communicating probabilities ONGOING 75%

Subtask 3: Develop evidence browser and explanation generator for TIC-MIL ONGOING 70% COMPLETE

Subtask 4: Continued development and refinement of CDS tool ONGOING 50% COMPLETE

Milestone #9 Prototype tool for clinical use of TIC-MIL and LIMB-MIL

Year 3 progress. Specific aim 3 and 4 are inherently co-dependent and related in that their respective outputs and findings of one will influence the research approach for the other. Two areas dominated this years progress: firstly a systematic review, conducted to refine the methodology by which user testing would be conducted, then a pilot prospective use-ability study of clinical decision-makers exposed to a mock-up of the prototype decision support system. The full study, for which Human Research Authority approval has been gained, is also in progress.

Systematic review. Prior to designing the prospective usability study, a systematic review was conducted to understand what methods are employed when testing the usability of mobile CDSS in emergency situations. This study was registered with PROSPERO: Jared Wohlgemut, Erhan Pisirir. Usability of mobile clinical decision support systems designed for clinicians treating patients experiencing medical emergencies: a systematic review. PROSPERO 2021 CRD42021292014

Material & Methods. A systematic review of usability assessment of mobile application CDSS in healthcare emergencies was conducted, utilising Pubmed/Medline, Embase, Scopus, and IEEE Xplore. After de-duplication, each title and abstract was screened by two reviewers. Screening-positive studies were full-text reviewed, along with references of selected studies. Risk of bias was assessed using the modified Downs and Black checklist. Quantitative data was descriptively analysed.

Results Of 974 articles from the initial search result, 23 studies were included. The most frequent usability evaluation criteria were efficiency and usefulness (both in 15 (65%) studies), followed by errors (61%), satisfaction (57%), learnability (48%), effectiveness (39%), and memorability (9%). Usability methods included questionnaires (90%), user trials (74%), interviews (26%) and heuristic evaluations (13%). CDSS inputs consisted of buttons/checkboxes (78%), numbers/text (35%), image inputs (9%), or automatic monitor and physical input (9%). CDSS outputs comprised a recommendation (78%), treatment (26%), or score, risk level or likelihood of diagnosis (26%). Of the five studies which measured the system usability scale, all achieved acceptable usability scores (>67), while studies evaluating technology acceptance model achieved mixed results.

Conclusions. Usability evaluations of mobile application CDSS differ widely in methodological design. Adherence to standardised usability definitions and methodologies is warranted to improve generalisability and permit comparison between systems.

Figures and Tables.

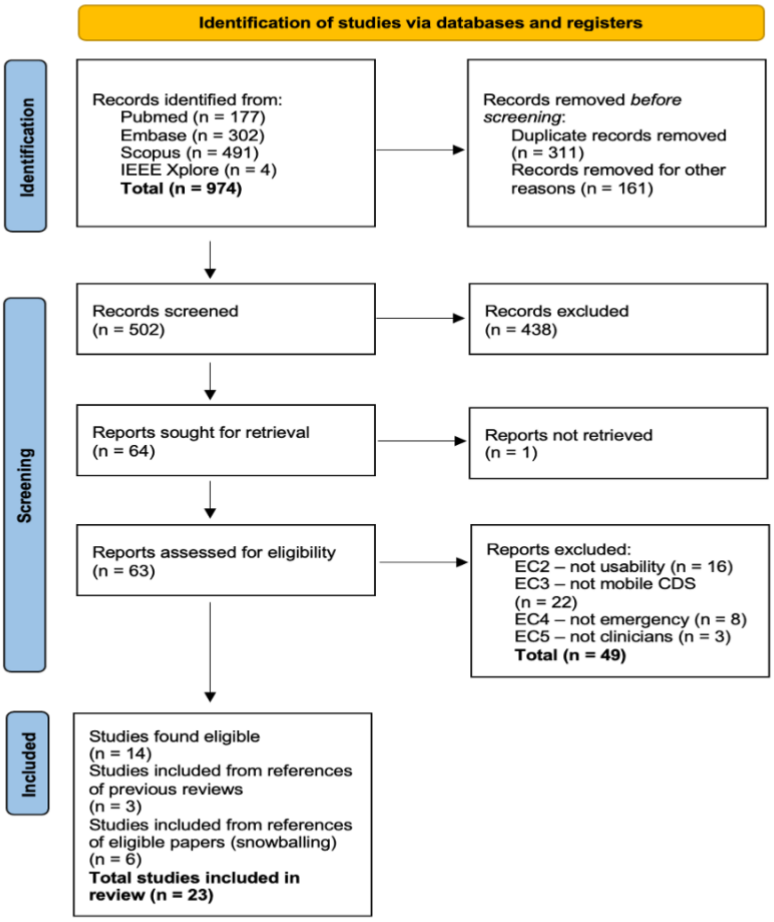


Fig 15- PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. After systematically searching 4 databases, 974 titles were retrieved. After duplicates were removed, and those which did not represent full peer-reviewed publications, there were 502 records. Titles and abstracts were reviewed, and 64 remained for full-text review (though on was not accessible). Of these, 49 were excluded based on exclusion criteria, and 14 were included. References of included studies and excluded reviews were included, adding 9. A total of 23 papers were included in the review.

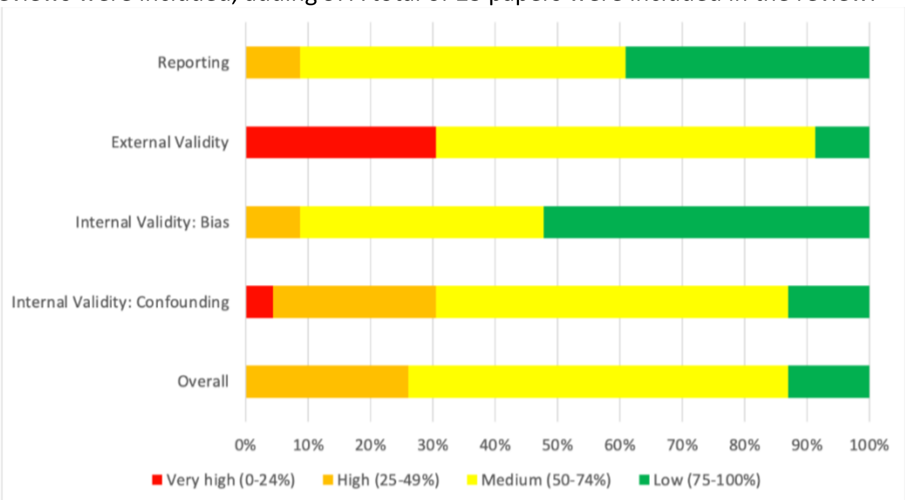


Fig 16. Quality assessment of included studies, using the modified Downs and Black criteria. Overall, most studies had a medium risk of bias, though 25% had a high risk of bias, and 12% had a low risk of bias. Overall, no study had a very high risk of bias.

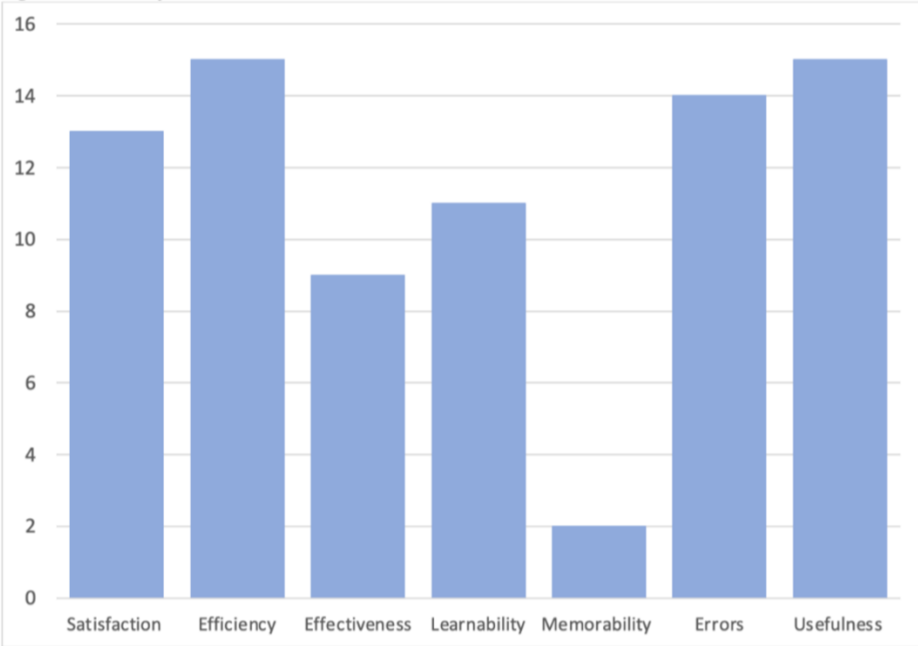


Fig 17. Usability evaluation criteria of included studies. The most frequent usability evaluation criteria were efficiency and usefulness (both in 15 (65%) studies), followed by errors (61%), satisfaction (57%), learnability (48%), effectiveness (39%), and memorability (9%).

Prospective Usability Study: Informed by the systematic review of usability evaluation above, we have designed a prospective study to evaluate the usability of our clinical decision support system (CDSS) with invited trauma clinicians. The usability study protocol now has IRAS/HRA approval (similar to IRB), and NHS Capacity and Capability approval to commence recruitment.

Because of the novelty of the methodology, after designing the study we conducted a planned Pilot study with study team members, in order to 1) to ensure the study methodology was fit for purpose, and 2) to test the usability of the CDSS, and make iterative improvements prior to prospective study. This allowed early identification and correction of issues with our CDSS. Five members of the study team participated in the pilot study, which was followed by a focus group discussion. The feedback from these have been used to improve the CDSS before moving on to usability study sessions with participants. An abstract is shown below, of this Pilot Usability study, which has been prepared for ECTES 2023 (European Congress of Trauma and Emergency Surgery, Ljubljana, Slovenia).

Material & Methods. This was a prospective, mixed-methods, pilot usability study. Participants tested the CDSS by inputting information into a user interface, generating a risk prediction output. During user testing, “think aloud” methodology captured usability issues, categorised according to Nielsen’s Heuristics (NH), and graded for severity (1=cosmetic, 2=minor, 3=major, 4=catastrophic). A validated questionnaire was administered based on the System Usability Scale (SUS), Technology Acceptance Model (TAM), and nonadoption, abandonment, scale-up, sustainability and spread (NASSS) adoption framework.

Results: Five participants highlighted 18 study design problems: 13 user testing, 5 questionnaire-related. Usability issues were identified within 9/10 NH categories (Fig 1). There were 40 input- and 7 output-related usability issues (average severity 2.1/4 and 1.4/4, respectively). Most common was ‘consistency and standards’ (n=14). The CDSS achieved an SUS score of 60 (Fig 2). Sixty percent of users said the CDSS would improve their clinical performance and would be useful in their clinical practice, and 40% said it would enhance their clinical effectiveness. Potential barriers related to the adopters, organisation and external environment were most prevalent.

Conclusions: This pilot identified several issues with usability study design and the AI CDSS. After iterative improvements, a prospective study will evaluate the AI CDSS' usability, clinician trust, and impact on decision-making.

Figures and tables

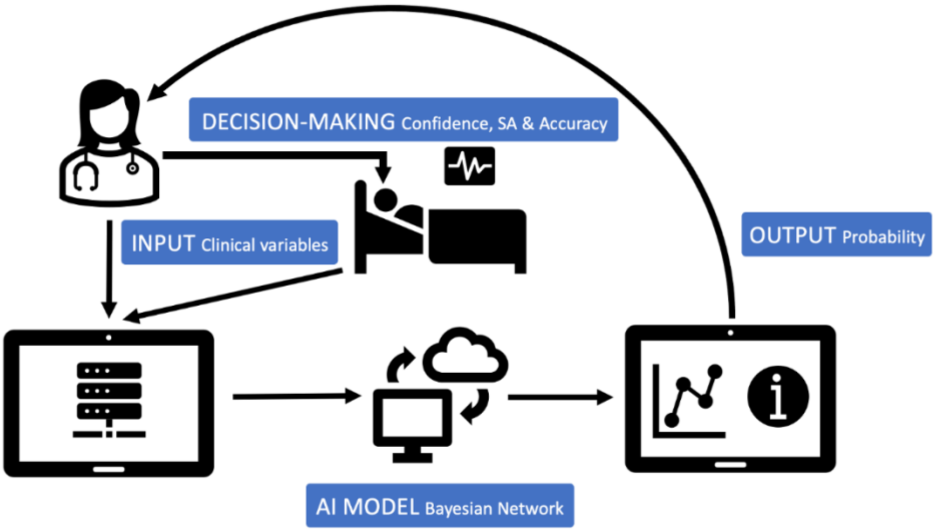


Fig 18. Conceptual model of Prototype CDSS. The INPUT will consist of variables entered from the clinician, and entered automatically from patient devices (monitors, etc), into the user interface of a CDSS. This information will be processed by the AI MODEL (Bayesian Network), to produce a probability, and present this in the OUTPUT, along with an explanation of the model's reasoning, back to the clinician via the same user interface. This output then assists the clinician in DECISION-MAKING, including their confidence, situational awareness and decision accuracy.

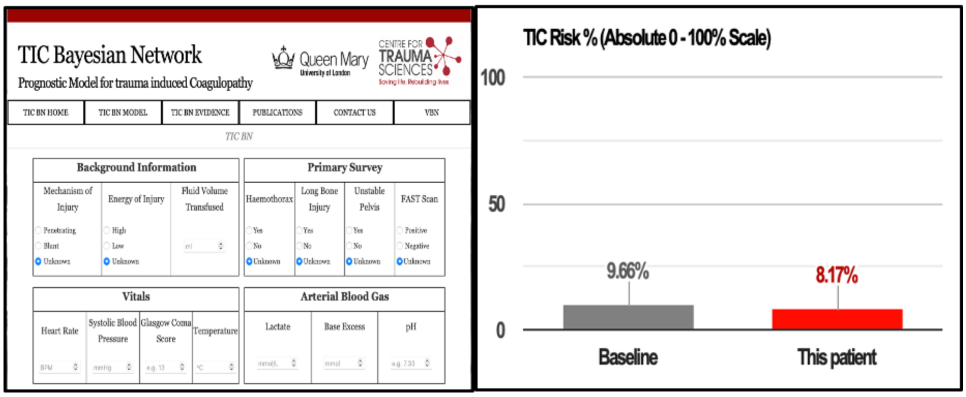


Fig 19. Old prototype CDSS (input and output). This is the version of the BN model input (on left), which currently exists on www.traumamodels.com, and the corresponding output (on right).

Fig 20. New Prototype CDSS (input and output), which was pilot-tested: This version of the CDSS was designed to include the inputs which match the updated BN model (TIC-MIL) in an intuitive user-friendly way (on left), and the corresponding output which provided the output as an absolute value, as well as within a risk stratification (from very low to very high risk of TIC; from green to red, on right). This is the version which was pilot-tested with study team members.

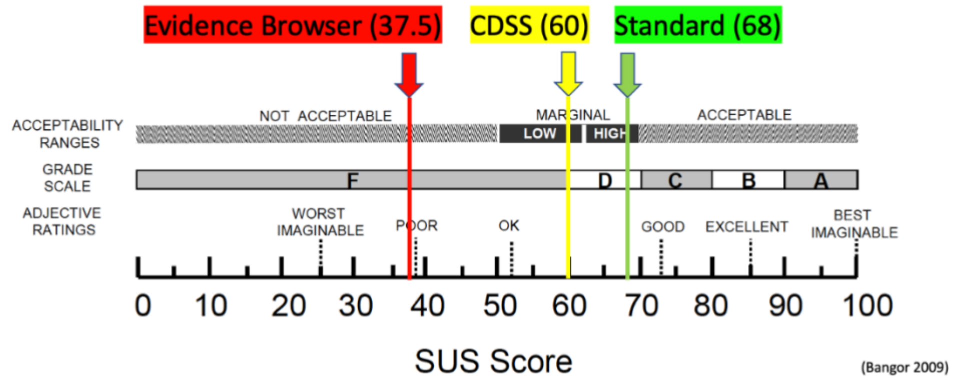


Fig 21. System Usability Scale (SUS) result of CDSS. The CDSS shown in Fig 3 was evaluated in the pilot test, and these are the results according to the 10-question SUS, presented as acceptability ranges, grade scale and adjective ratings, and number out of 100. Our previous evidence browser was not acceptable in its current form, achieving a score of 37.5 ("Poor" usability). Our CDSS achieved a marginal, or "OK" usability score. However the standard across industries is considered 68 or above. Part of the pilot-testing was also evaluating the mental effort required to use the tool and the specific usability issues which need addressing.

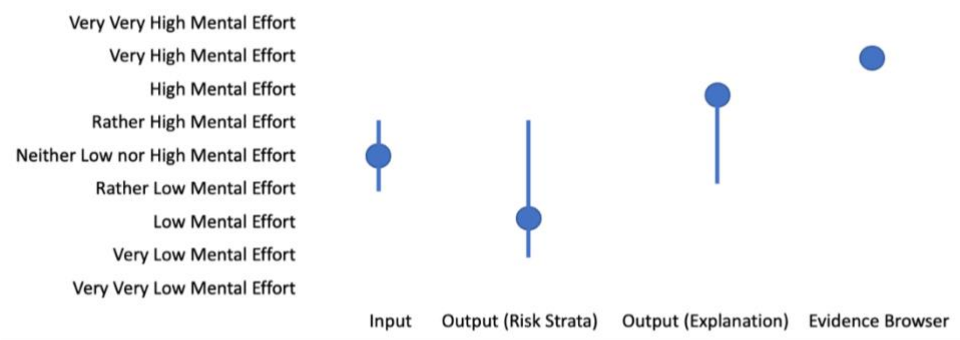


Fig 22. Mental effort required to use CDSS, according to Paas scale. From the pilot testing, the median mental effort required to use the CDSS input was "neither low nor high", to use the output (risk strata) was "low", to use the output (explanation) was "high", and to use the evidence browser was "very high". Despite low numbers, this provides an indication as to which aspects of our CDSS which we need to concentrate on, to reduce the mental effort required to use the system, and thus improve usability.

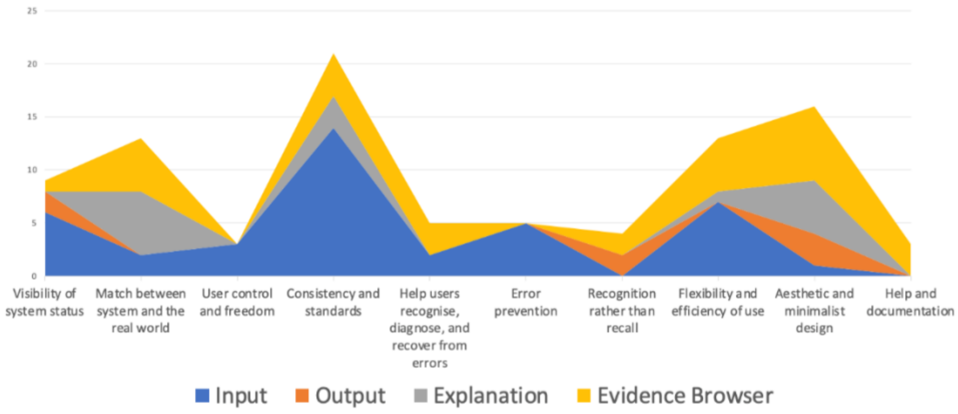


Fig 23. Nielsen's Heuristics of Usability design. Nielsen published 10 usability design heuristics, which have become a widely-used way to document and categorise barriers to the use of a system. This figure shows that the most frequent usability issues were regarding "consistency and standards" in the input, both "aesthetic and minimalist design" in the output, "match between system and the real world" in the explanation, and "aesthetic and minimalist design" in the evidence browser. Each issue was considered, and grouped together into themes on a "Miro" board. We then addressed each issue, and presented an updated version of the CDSS prototype to our study team, during a focus group discussion. Final tweaks from that focus group resulted in the final CDSS prototype which we will use for prospective usability testing, shown in Fig 7.

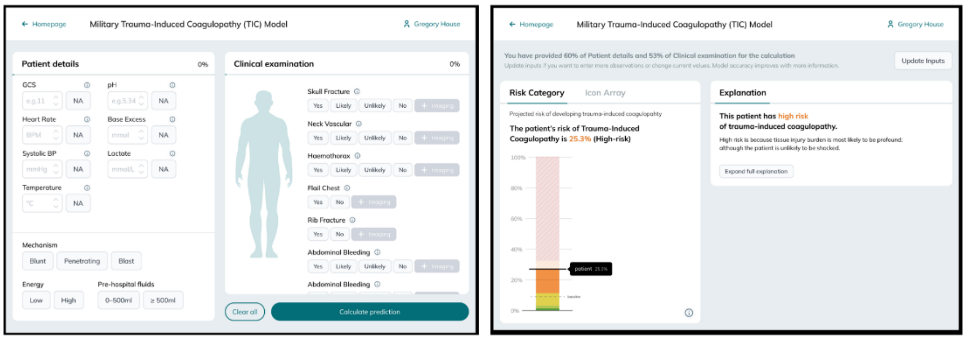


Fig 24. New prototype CDSS (input and output), improvements made after pilot-testing. This is the latest CDSS input (left) and output (right). It has been informed by usability testing and iterative improvements made.

The CDSS shown in Fig 24 will be tested with invited participants, to test its usability, usefulness, credibility, as well as explore potential barriers to adoption. Study procedures include user testing of the CDSS, followed by participating in a validated questionnaire (based on System Usability Scale, Technology Acceptance Model (TAM), and nonadoption, abandonment, scale-up, sustainability and spread (NASSS) adoption framework), and finally taking part in a semi-structured interview, to explore usability, decision-making, situational awareness, and barriers to CDSS adoption. This study is live, and has recruited 6/20 participants (30%), with 2/20 already interviewed (10%).

Specific Aim 5 – Develop and validate a prognostic model for blood product requirements

Major Task 2 – Develop a prognostic model for blood product requirements INITIATED, 25% COMPLETE

Subtask 1: Literature review and develop structure of Bayesian Network. [Will include a systematic review of predictors of blood transfusion requirements following injury] ONGOING 20% COMPLETE

Subtask 2: Construct a development dataset from UK military data sources COMPLETE

Mst #10 Complete development dataset for the blood requirement prognostic model

Subtask 3: Learn parameters of the blood requirement Bayesian Network ONGOING 30% COMPLETE

Subtask 4: Cross-validation and model refinement NOT INITIATED

Major Task 3 – External validation of prognostic performance

Subtask 1: Construct external validation dataset from US military data sources NOT INITIATED

Mst #11 Complete validation dataset for the blood requirement prognostic model

Subtask 2: External validation of prognostic performance NOT INITIATED

Milestone #12 Validated prognostic model for blood product requirements

Year 3 update. The aim of this work is to develop a model to predict rates of blood product consumption following trauma. A literature review of transfusion prediction models and predictive factors is underway, as well as analysis of our military data on blood product usage. A provisional review of the literature identifies multiple prediction models from both civilian and military fields, primarily designed to predict any transfusion, or massive transfusion (MT), rather than actual volumes of blood products required. Current models heavily utilise physiological variables, with almost all incorporating measures of systolic blood pressure and heart rate. The inclusion of anatomical injury variables tends to be restricted to the presence of pelvic or femur fracture. Recent machine learning models in this area have demonstrated excellent performance, but rely on variables that are not available at initial assessment, limiting practical application as predictions of this nature are likely of most use prior to embarking on surgery

.

Our intent remains to develop a Bayesian network model that incorporates early clinical information from both anatomical injury and physiological variables to predict transfusion volumes.

Data source and initial survey. Within this last reporting period, we have acquired additional data from the UK MoD. This is a rich and highly valuable source which comprises more than 65,000 units of blood products that were assigned to casualties treated at Camp Bastion, including UK military as well as other combatants and civilians. These units are time stamped and co-ordinated with time of injury and arrival to the Role 3 facility. Cases also have mortality data, including time to death. Focussing on PRBC delivery, we have data for 20,827 accurately time-labelled units administered to 2187 casualties within 24 hours of injury. It has been matched up with demographic, injury and intervention data held within our main JTTR dataset of UK combat casualties.

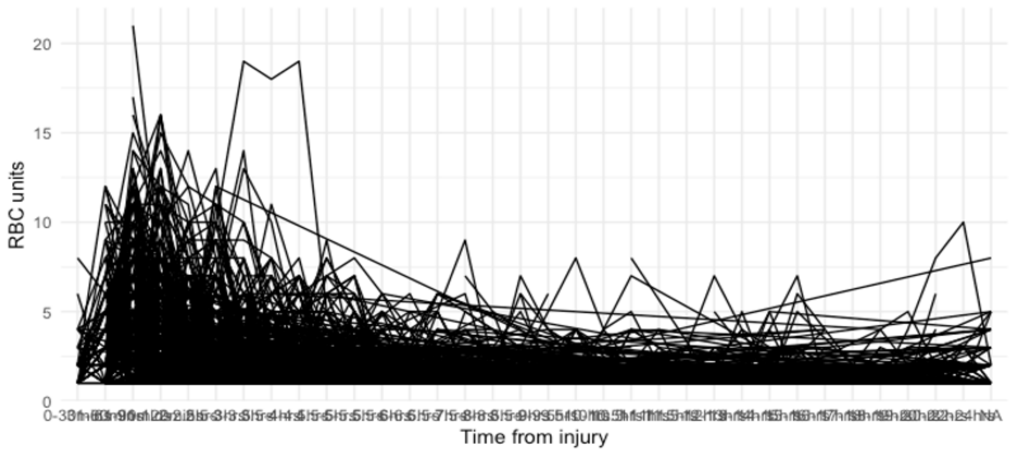


Fig 25. Draft scatter plot of units transfuse (vertical axis) against time (horizontal axis).

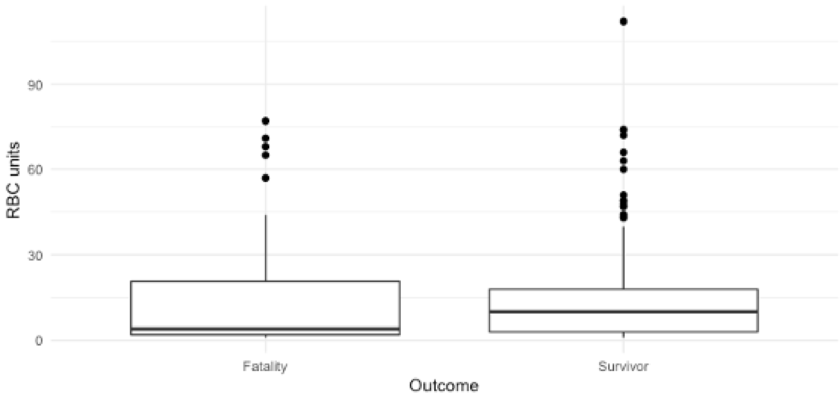


Fig 26. Box plot of transfusion of red blood cells. Usage per patient ranges from 1 - 112 units, with median usage 10 units for survivors and 4 units for fatalities. This is in the context of all casualties in this dataset having received at least 1 unit of blood.

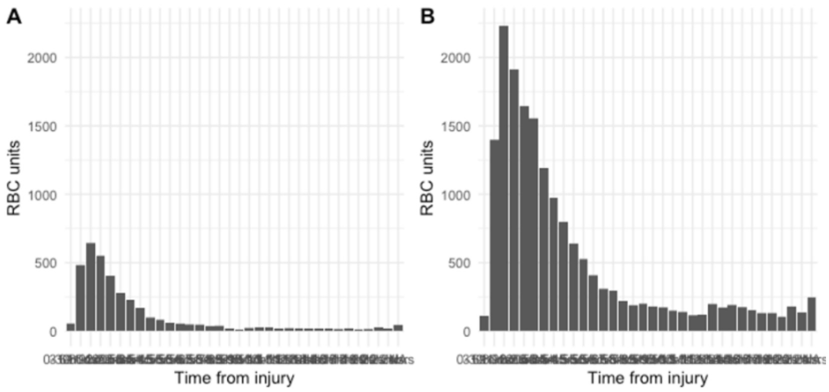


Fig 27. Histograms of transfusion with respect to time dichotomized by outcome. Resource usage by time is considered, there is substantial difference between the fatality (bar chart A) and survivor (bar chart B) cohorts. Fatalities make up 16.96% of all casualties in this dataset.

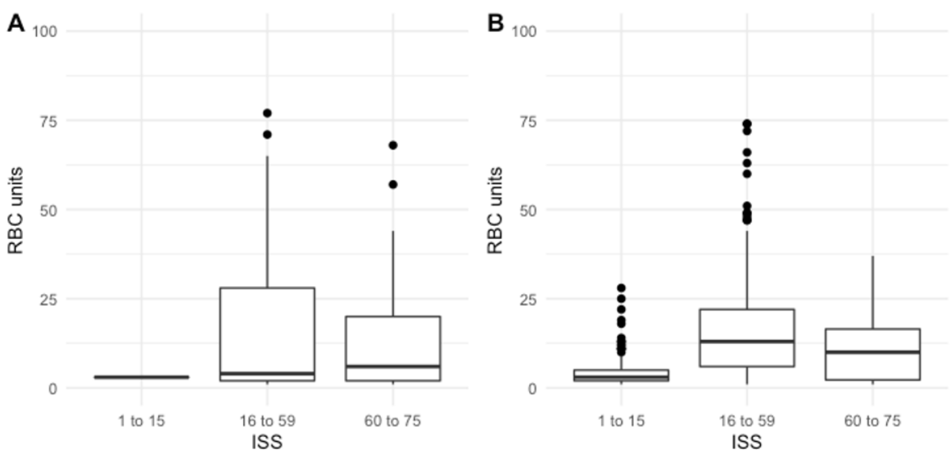
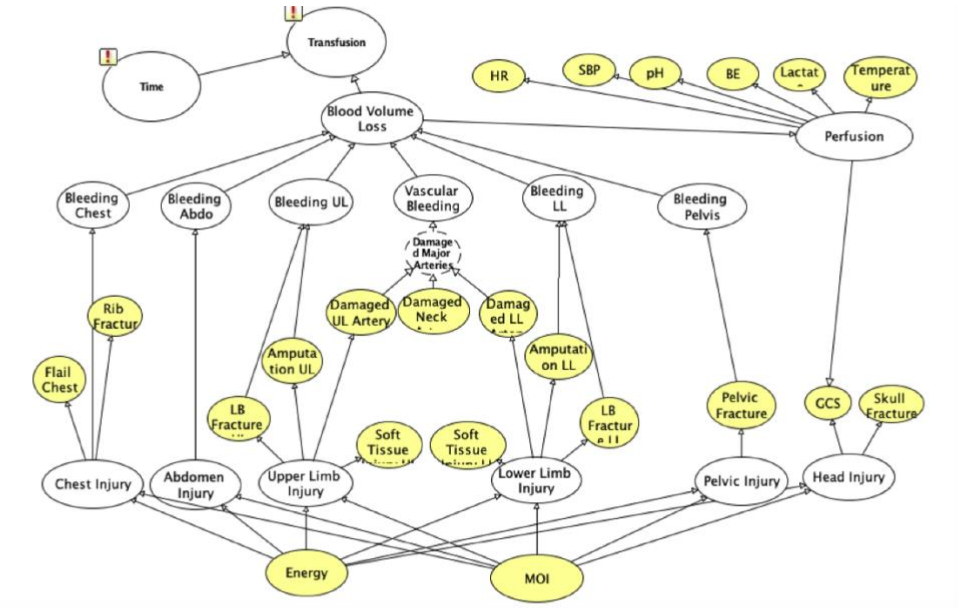


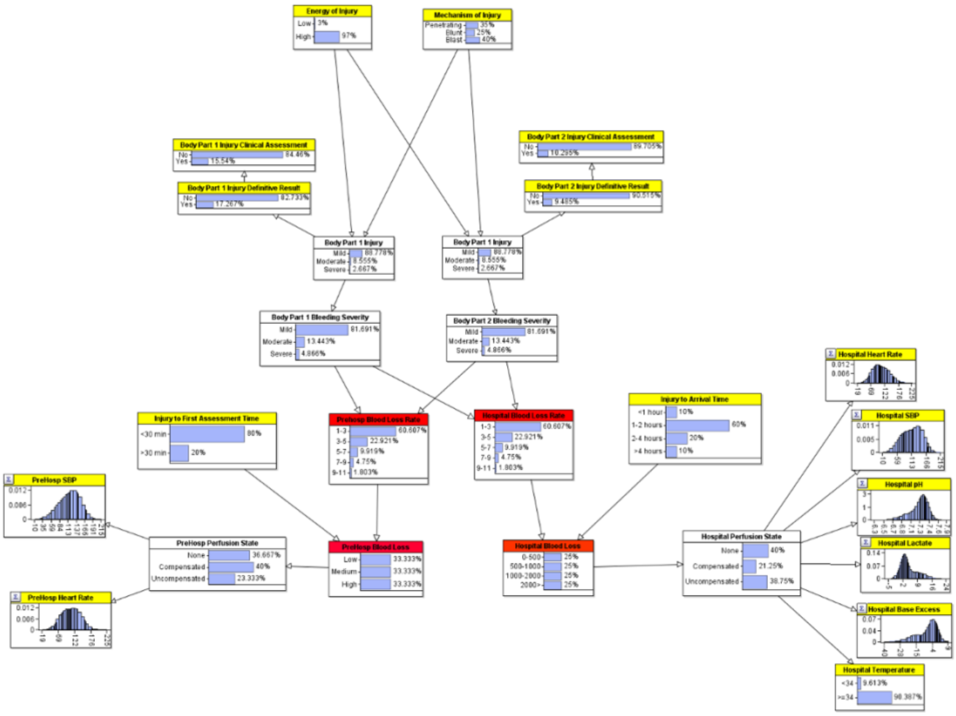
Fig 28. There is higher usage of blood products per patient for survivors (Boxplot B) compared to fatalities (Boxplot A) across all ISS groups. However, there is no statistically significant increase in PRBC usage when all casualties with ISS >15 are compared (p=0.40).

Planned work. Cluster analysis, using latent class modelling and hierarchical clustering methods, is ongoing to identify distinct patterns of transfusion over time in the initial resuscitation phase and identify associations with outcome, mechanism, injury pattern, or other variables. This cluster analysis work will be used to further inform our model structure development.

At this time, two approaches are being pursed to develop an accurate transfusion prediction model. Firstly, the extension of the TIC-MIL model to predict categorical blood transfusion volumes. The TIC-MIL model incorporates anatomical injury variables that correspond with a measure of estimated blood loss, which was established using a multiple expert consensus and categorised as mild, moderate, severe, or profound per injury. It is therefore hypothesised that the combination of this with time since injury and measures of perfusion state (physiological variables such as HR, SBP, pH, base excess, and lactate) should accurately predict blood product usage. This model structure is being utilised with the recently acquired UK military transfusion data, for training and internal validation. The structure is seen below in Figure 29:



The second approach is to develop a two-stage model that predicts both the rate of blood loss, from very early (pre-hospital/R2&3) clinical information, and total volume loss, utilising information gained on hospital/R2&3 assessment. This model outcome would then be translatable to a transfusion volume prediction. This will again be trained and internally validated using UK military data. The initial model structure is shown below (figure 30):



Specific Aim 6 – Develop and validate a prognostic model for mortality in injured military personnel (SURVIVAL-MIL).

Major Task 4: Develop prognostic model for mortality in injured military personnel
ONGOING 70% COMPLETE
we
Subtask 1: Literature review and develop structure of Bayesian Network. [Will include a systematic review of prognostic factors for mortality following injury] COMPLETE

Subtask 2: Construct a development dataset from UK military data sources. COMPLETE

Mst #13 Complete development dataset for the SURVIVAL-MIL prognostic model

Subtask 3: Learn parameters of the SURVIVAL-MIL Bayesian Network COMPLETE

Subtask 4: Cross-validation and model refinement ONGOING 70% COMPLETE

Development of the SURVIVAL-MIL model

Year 3 progress. Development of the mortality model has two strands - the development of a model structure, via single vs combined structure comparison, and the recalibration of parameters from a civilian to a military context. SURVIVAL-MIL is designed to incorporate what were considered the main causes of early traumatic deaths - early, in this case, refers to the exclusion of secondary complications such as emboli, thrombosis, and infection. The main causes addressed in this model are coagulopathy, organ failure, and traumatic brain injury. Model structure development was based around the concept of developing individual models for predicting (or in the case of TBI, identifying) each condition. This required the development of models for organ failure and TBI.

Organ failure model. Rather than attempt to predict any or all organ failures, a surrogate organ was chosen to represent all organ failure. This should be a clearly identifiable, classifiable, organ failure; ideally that presents early. Renal failure was therefore chosen, as a failure with clear diagnostic criteria, as well as being an organ failure that occurs relatively early. The Trauma-induced Acute Kidney Injury (TAKI) model was developed using the same hybrid methodology as the TIC models. The development and performance of the TAKI model in civilian trauma can be seen in the abstract below, presented at AAST (Chicago, September 2022).

Introduction: Trauma-induced Acute Kidney Injury (TAKI) occurs in ~20% of trauma patients admitted to Intensive Care (ITU). Accurate prediction of which patients will develop TAKI requiring intervention such as renal replacement therapy (RRT) can ensure early resource allocation and treatment, minimizing mortality. We aimed to develop an artificial intelligence (AI) risk prediction model (TAKI-BN) for TAKI using information available at two time points: on first assessment in the emergency department (ED) and on admission to ITU.

Methods: Training and validation data was from the Activation of Coagulation and Inflammation in Trauma study, including adult trauma patients admitted to a UK Major Trauma Centre <2 hours since injury. Patients with length of stay ≤1 day or no serum creatinine measurement were excluded. The algorithm is a Bayesian Network (BN). BN structure was developed from literature and expert knowledge, to include known variables that influence TAKI risk and reflect known causal relations. Parameters were learned from data. Model outcome was risk of TAKI, classified as worst AKI state within first 3 days. KDIGO states 0, 1, 2&3 were classified as None, Mild, Severe. 10-fold cross-validation was undertaken and performance assessed through discrimination (Area under the receiver operator curve (AUROC)) and calibration (Slope and intercept) for binary outcome (None vs Mild/Severe), and accuracy for both binary and categorical outcomes.

Results: Dataset comprised 1234 patients with median age 36, 81% male, median Injury Severity Score 17, 20% penetrating mechanism, mortality 11%. Overall, 32% developed AKI within 3 days of admission, of which 68% were mild and 32% severe. Mortality was 7% and 33% respectively. Internal validation demonstrated excellent performance at ED time point (AUROC 0.93, calibration slope 1.034 and intercept -0.018, accuracy 0.87), as well as excellent at the ITU time point (AUROC 0.93, slope 1.020, intercept -0.005, accuracy 0.88).

Conclusions: An individual patient's risk of TAKI can be reliably predicted from information available at initial assessment as well as following resuscitation. This information can be used to allocate treatment and resources to those who need it most.

Figures and tables

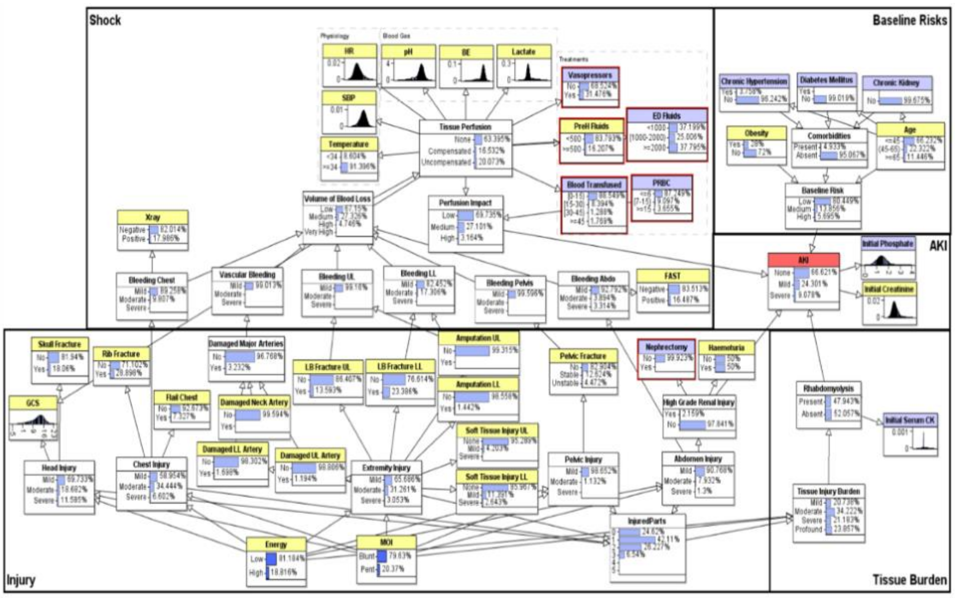


Fig 31. The Bayesian Network underpinning the TAKI model.

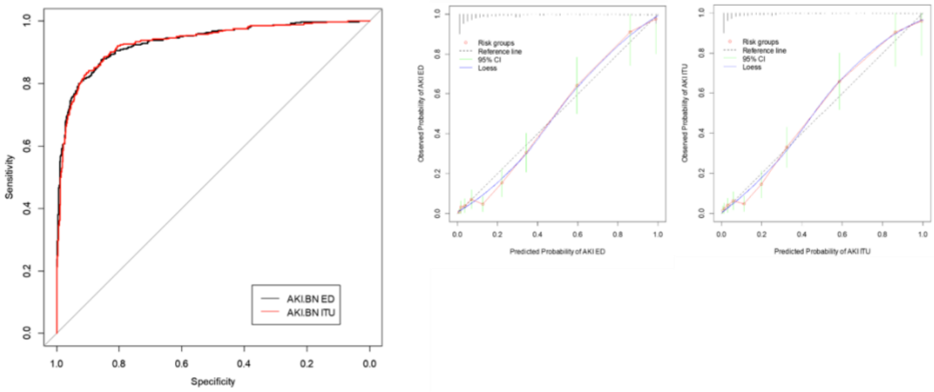


Fig 32. Area under the receiver operator curve (left) for prediction model performance base on variables harvested from Emergency Department (black) and Intensive Care (red) patient data (left hand graft), with calibration curves (middle and right hand graft).

Recalibration of TAKI for military use is underway, using UK JTTR data. The use of the TIC-MIL structure segments for incorporating mechanism, anatomical injuries, and physiological variables mean that updating of the model structure for military use is not anticipated to be necessary.

Traumatic Brain Injury model. TBI posed a unique challenge in predictive modelling. The condition has already occurred at the point of model use, but the key is in identifying the likelihood of its existence without the use of imaging technology. This would enable the mortality model to be used in the pre-hospital/R3 and pre-CT phases of care. Many predictive models exist for prognosis after TBI, particularly relating to mortality and disability. Most of these are based upon CT findings, such as the Rotterdam or Marshall classifications.

Methods. There were few variables that could be readily utilised for this model, that fit the criteria of being available pre-hospital and pre-imaging, whilst being strongly associated with the presence of TBI. The model structure was developed with the variables HR, SBP, GCS, and presence of skull fracture. The model outcome was the presence of TBI. However, there are multiple ways to define this, with no agreed standard. A combination of AIS and GCS was used. The case was classified as having 'No TBI' if the Head AIS (separated manually from neck injuries that are classified together as HNAIS) was less than 3. Where the AIS was ≥3, GCS was used to grade the severity, with 14-15 as mild, 9-13 as moderate, and ≤8 as severe.

The structuring of the combination of HR and SBP was considered with several approaches. Initially, they were both utilised as direct parents of TBI, as continuous variables. However, this created a nodal probability table of almost unlimited combinations. Combining them in a rules based fashion was therefore considered, taking into account the concepts of shock index (HR/SBP) and reverse shock index (SBP/HR).

The purpose of HR and SBP as variables in this model was to differentiate between the cases with low GCS due to hypovolaemic shock, and those with low GCS secondary to brain injury. Utilising SI and rSI was intended to define those patterns.

Classification was therefore proposed with 3 categories:
Positive SI: HR > SBP
Neutral: HR < SBP and HR > 80 or SBP <120
Negative/Reverse: HR < SBP, where it does not meet the criteria for Neutral
The model structure is seen below:

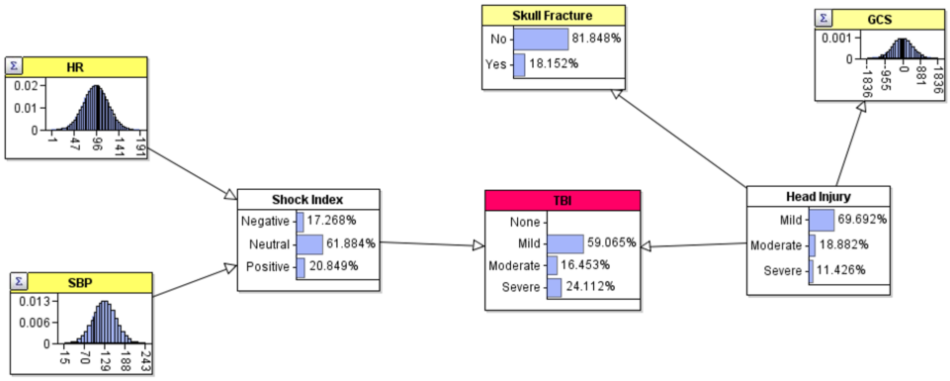


Fig 33. Draft model for prediction of Traumatic Brain Injury related mortality. This model is in the process of training and validation, using a dataset of 1234 civilian ACIT cases. It will then be tested with military UK JTTR data.

Overall SURVIVAL-MIL model. There were several approaches considered in the development of the overall (combined) mortality model. Firstly, (Approach A) that each sub-model (TIC, TAKI, TBI) could be used independently, the risks of each condition calculated, and these used to calculate the overall risk of mortality. This would require more work to be focused outwith the model itself, instead being more interface development.

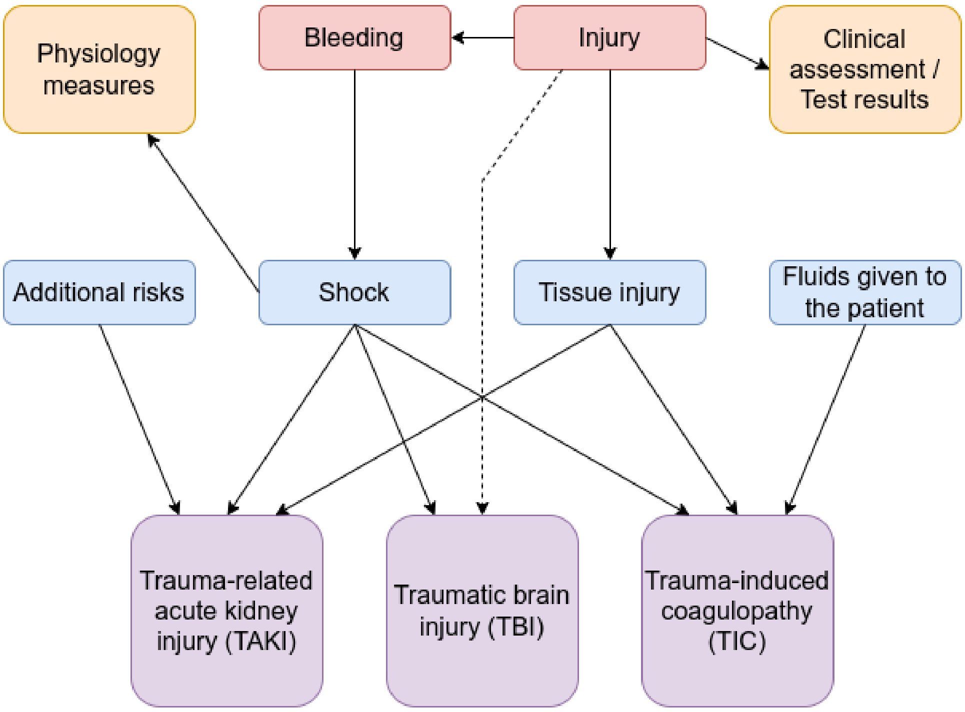
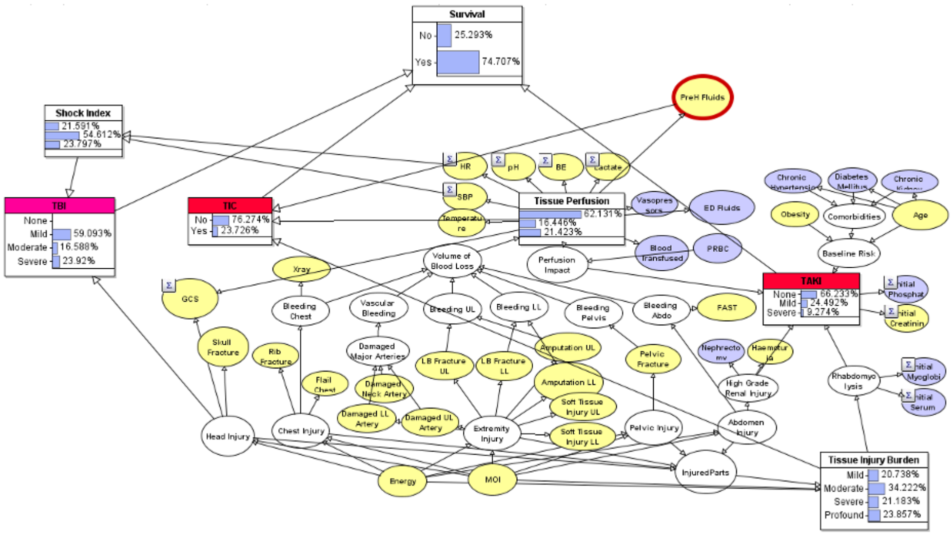


Fig 34. Schematic of proposed Approach A to field a mortality prediction that discriminates between death due to organ failure, brain injury or coagulopathy.

Alternatively, the three models could be combined to produce a single prediction (below, Figure 35). All three utilise the same variables; that is, all variables of the TBI model are found in both TIC-MIL and TAKI, and TIC-MIL contains only one variable not found in TAKI. However, the parameters within each model are different. Initial training of a combined structure proved difficult, as it exceeded the computational power of a desktop computer. This is being overcome with the use of cloud-based server capability.



Whilst it was identified that TBI is an important and significant contributor to trauma mortality, the issue of it being an identification model rather than a truly predictive model has led to the reconsideration of its inclusion within the mortality model.

We are currently assessing the performance of SURVIVAL-MIL including the TBI structure, compared to without the latter. The concept of head injury remains captured within the injury variables, and the physiology variables are also represented. It is hypothesised that TBI does not require a separate predictive structure to accurately predict overall mortality, as it is sufficiently well represented elsewhere in the combined TIC-MIL+TAKI structure.

Once validation of the mortality model is complete with civilian data, we will recalibrate it for military use. The structure has been based upon the military version of TIC, so the structure does not require updating for this purpose. The recalibration will be performed using data from the UK JTTR. If the non-TBI model performance is sufficient, the combination of TIC-MIL and the recalibrated TAKI model will not require further recalibration for military use.4

Major Task 5 – External validation of SURVIVAL-MIL prognostic performance NOT INITIATED

Subtask 1: Construct external validation dataset from US military data sources NOT INITIATED

Milestone #14 Complete validation dataset for the SURVIVAL-MIL prognostic model

Subtask 2: External validation of SURVIVAL-MIL prognostic performance NOT INITIATED

Mst #15 Validated prognostic model for military trauma mortality (SURVIVAL-MIL)

Year 3 progress. This work is dependent upon availability of US JTTS data and the product of major task 4.

Specific Aim 7: Develop and validate prognostic model that can quantify patient-specific risks related to ‘right place’ and ‘right time’ care (OVERFLIGHT-MIL) NOT INITIATED

Major Task 1 – Develop the Bayesian Network NOT INITIATED

Subtask 1: Literature review and develop structure of Bayesian Network Initiated, complete in Year 4
Subtask 2: Construct a development dataset from UK military data sources. NOT INITIATED

Milestone #16 Complete development dataset for the OVERFLIGHT-MIL

Subtask 3: Learn parameters of the OVERFLIGHT-MIL prognostic model NOT INITIATED

Subtask 4: Cross-validation and model refinement A NOT INITIATED

Major Task 2: External validation of OVERFLIGHT-MIL NOT INITIATED

Subtask 1: Construct external validation dataset from US military data sources NOT INITIATED

Milestone #17 Complete validation dataset for the OVERFLIGHT-MIL

Subtask 2: External validation of OVERFLIGHT-MIL prognostic performance NOT INITIATED

Milestone #18 Validated OVERFLIGHT-MIL prognostic model

Year 3 progress. This work is dependent on the results of Specific Aim 6 and has yet to be initiated. A scoping framework is as follows:

Overflight-MIL is a completely different model than the others developed by our research group. The purpose of the model is to help decision-makers with advanced triage on the battlefield. More specifically, to help decision-makers decide whether to bypass, or 'fly over' less equipped facilities in order to convey an injured casualty to an appropriate care facility for their needs. The premise is that, in order to survive, an injured patient's needs should be matched to available resources to treat those needs. A patient who is minorly injured may only need battlefield care (Role 1; R1), and the need may not be urgent. A patient who is severely injured may require advanced care provided at R2, R3 or R4. However, especially for severely injured patients, there may be a time frame within which certain aspects of care are required, depending on the nature of injuries. Equally, whilst time to care is important, so is the place of care. In some contexts it may be better to transfer a patient to a more replete and stable setting where there are sophisticated imaging, surgical and post-operative facilities, even if this means a longer time-to-definitive treatment. These triage and transport decisions become even more challenging in the context of multiple or mass casualty incidents. If there was a method to predict the treatment needs of a given patient, then this information could be matched to available resources and timings could be displayed to decision-makers, in order to make triage and transport decisions on the battlefield.

Levels (Roles) of Trauma and Injury Care

Current Levels (Roles) of Care	Function
Role I Battlefield Care to Battalion Aid Station	Initial level of care/immediate lifesaving measures. Emphasis on stabilizing casualty for evacuation to next level of care. Similar to civilian first responders. Also includes: Battlefield Care (Self-Aid/Buddy Aid, Combat Lifesaver and Combat Medic). Battalion Aid Station (far forward aid station with at least one physician available.
Role II Forward Surgical Team	Small, highly mobile, austere surgical team. Provides life-and-limb saving surgical care and typically the first level of surgery available. Limited capabilities, some laboratory, X-ray, mental health and dental services may be available.
Role III Combat Surgical Hospital Air Force Theater Hospital	High volume trauma center. Highest level of treatment within the area of military operations. Provides full range of surgical, medical, laboratory, and radiology capability. Care also includes dental, physical therapy, mental health, obstetrics/gynecology, and primary care services.
Role IV OCONUS Example: Landstuhl Regional Medical Center	Definitive medical and surgical care. Outside of area of military operations or combat, but not within CONUS. Stabilization point before evacuation to CONUS.
Role IV CONUS Walter Reed National Military Medical Center, Brook Army Medical Center	Definitive medical and surgical care OCONUS.

Table 2. Levels (Roles) of trauma injury care (Lane et al, The Afghan Theater: A Review of Military Medical Doctrine From 2008 to 2014. Military Medicine, 2017;182(3/4):32.2017)

Proposed Methods. Plans include a utility model, which first calculates patient needs (based on mechanism, physiology, injuries, blood tests, imaging results and some treatments (blood transfusion and haemorrhage control interventions). Outputs will depend on patient risks, available resources, and time from injury to arrival to each medical treatment facility (MTF). Table 2 outlines the patient inputs which are needed to inform the BN model risk predictions, which BN risk predictions will be calculated, the capabilities and transport distances of available medical facilities, and the hypothetical utility output for each facility, which is shown to the decision-maker. This allows a simple numerical (or multiple numerical) risk prediction to be calculated, whether the patient is conveyed to each available facility. An illustration of this is shown in Figure 1, with the decision support points relating to (in this example), what the likelihood of patient survival to that role/facility, and whether the capabilities of that facility are adequate to deal with all the patient's needs.

Table 3. Information which must be input, BN risks predicted, facility inputs, and hypothetical utility output.

Patient Inputs	BN Model Risk Prediction (=Need)	Facility Inputs (=Resource)	Hypothetical Utility Output (Matching Need to Resource)
- Mechanism - Physiology - Injuries - Treatments (blood transfusion, haemorrhage control interventions) - Blood results - Imaging results	- Trauma-induced coagulopathy - Lower limb viability - AKI - Transfusion requirement (over time) - Survival	- R1= transport distance, capabilities - R2= transport distance, capabilities - R3= transport distance, capabilities - R4= transport distance, capabilities	Given patient A has needs XYZ, if taken to facilities 1234, the likely outcome would be: R1- survival? meets pt needs? R2- survival? meets pt needs? R3- survival? meets pt needs? R4- survival? meets pt needs?

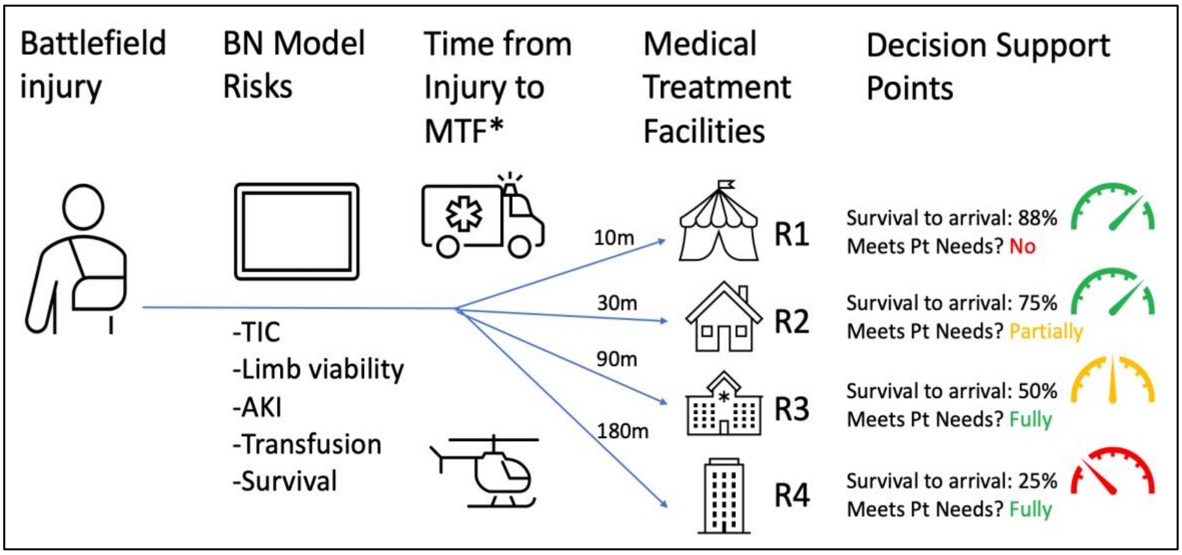


Fig 36. Illustration of the potential use of the Overflight clinical decision support system.

Specific Aim 8: Extend CDS tool from Aim 3 to include prognostic models developed in Aim 4, 5, and 7. 60% COMPLETE

Subtask 1: Determine most effective methods of presenting raw probability outputs to support precision medicine. [User simulation experiments and interviews] ONGOING 65% COMPLETE

Subtask 2: Develop user-interface to capture information and present probabilities ONGOING 65% COMPLETE

Subtasks 3-5: Develop the evidence browser and explanation generator for the blood product requirement, SURVIVAL-MIL, and OVERFLIGHT-MIL prognostic models. ONGOING 65% COMPLETE

Milestone #19 Functional evidence browsers for blood product requirement, SURVIVAL-MIL and OVERFLIGHT-MIL prognostic models.

Specific Aim 9 - Evaluate the clinical usability of the prototype CDS tool and validate the effect the tool has on the clinical credibility of the suite of prognostic models. ONGOING 60% COMPLETE

Subtask 1: Assess the clinical usability of the user-interface ONGOING 65% COMPLETE

Subtask 2: Assess clinical usability of evidence browser and explanation generator ONGOING 65% COMPLETE

Milestone #20 Prototype tool for clinical use of the suite of prognostic models

Subtask 3: Develop a protocol for a clinical impact study of the CDS tool COMPLETE

This aim is related and co-dependent on specific aims 3 and 4 and the research endeavours of 3, 4 and 9 are being pursued contiguously.

Subtask 3 is complete as reported in last year's report (with study protocol submitted as part of a CDMRP FY 2021 grant application RE: JW210420 - "Evaluating the Impact of an AI-Powered Clinical Decision Support Tool in the Management of Major Trauma". The research team incorporated the protocol in the grant application, which was recommended for Alternate Funding. The final status of the application was communicated to the PI on September 23 2022 (FY 2021 funds not available).

CONCLUSION

Year 3 of this project has seen satisfactory progress, with the Military TIC model complete (and awaiting validation on US data) and the Military Mortality model (integrating TIC, TAKI and potentially TBI) near complete (and awaiting internal UK validation and external US data). The work on the blood transfusion model is ready to commence now that the UK MoD blood transfusion data is ready for exploitation. Overflight model building can commence once the mortality work is complete. The prototype Interface for the CDSS is progressing very well, informed by the end-user testing which has completed pilot stage and is ready for the full study. Work has been presented in Oslo, Chicago and Florida. Workforce challenge has been mitigated through the granting of a no-cost extension (12 months). A decision on applying for a further no-cost extension will be made in early 2023.

Colonel Nigel RM Tai CBE MS FRCS

Principle Investigator