THOR Solstrand Hotel, Norway 19-21st May 2025

Artificial Intelligence: Bleeding

Max Marsden

Honorary Clinical Lecturer, Centre for Trauma Sciences, QMUL, UK

Senior Lecturer in Surgical Informatics, Academic Department of Military Surgery and Trauma







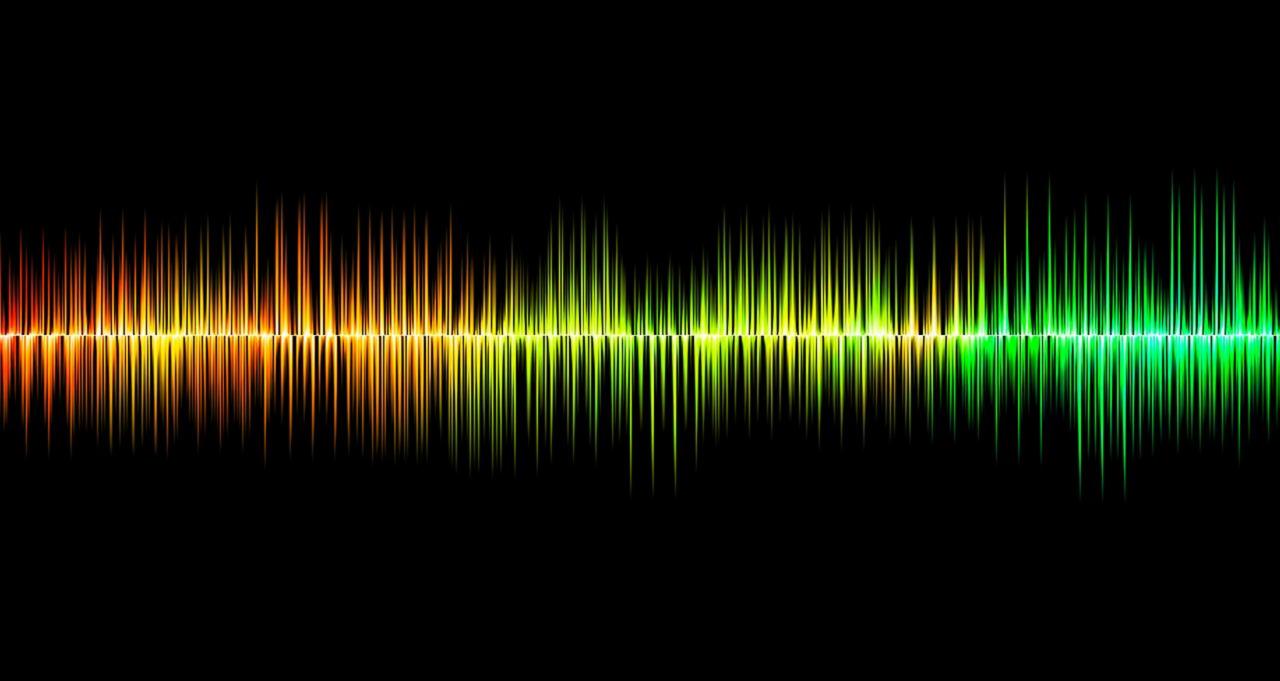


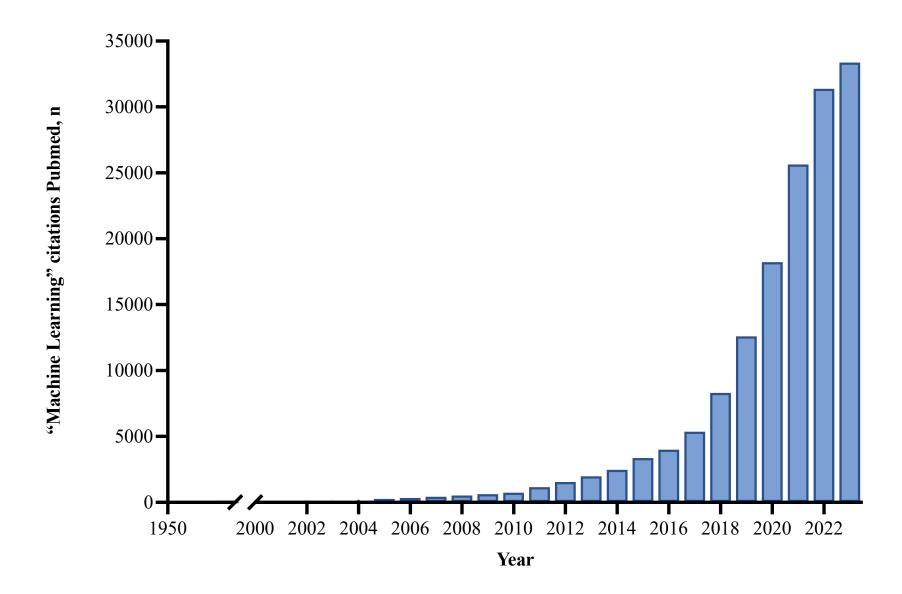


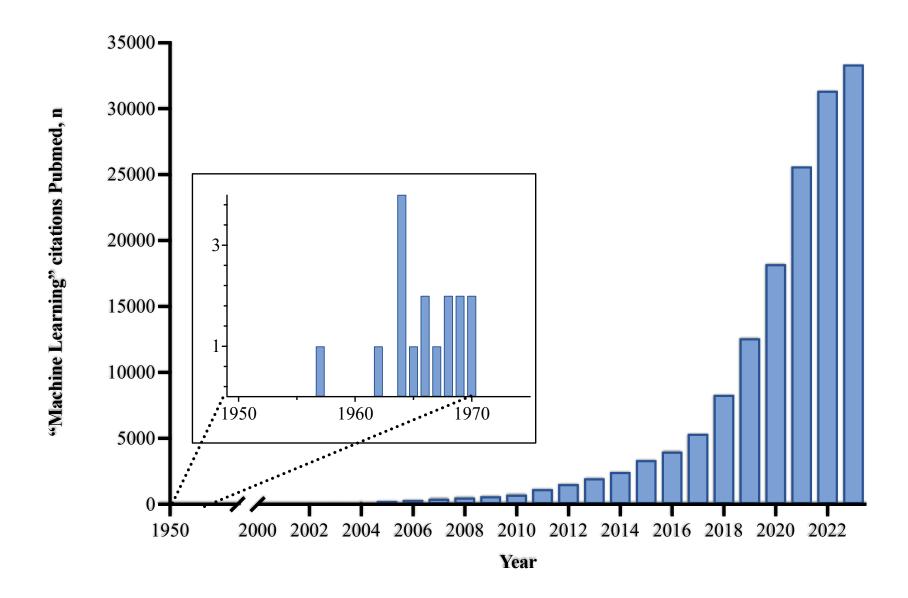












MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. The Imitation Game.

I PROPOSE to consider the question, 'Can machines think?'
This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively upambiguous words.



Workflow Enhancements	Diagnostics (e.g. image processing)	Decision Support	Predictive Analytics
Triage	Telemedicine	Robotics and Surgical Navigation	Monitoring & Alarms
Rehabilitation (wearables)	Psychological Aftercare	Research	Simulation & Training



Current knowledge and availability of machine learning across the spectrum of trauma science

Tobias Gauss^a, Zane Perkins^b and Thorsten Tjardes^c





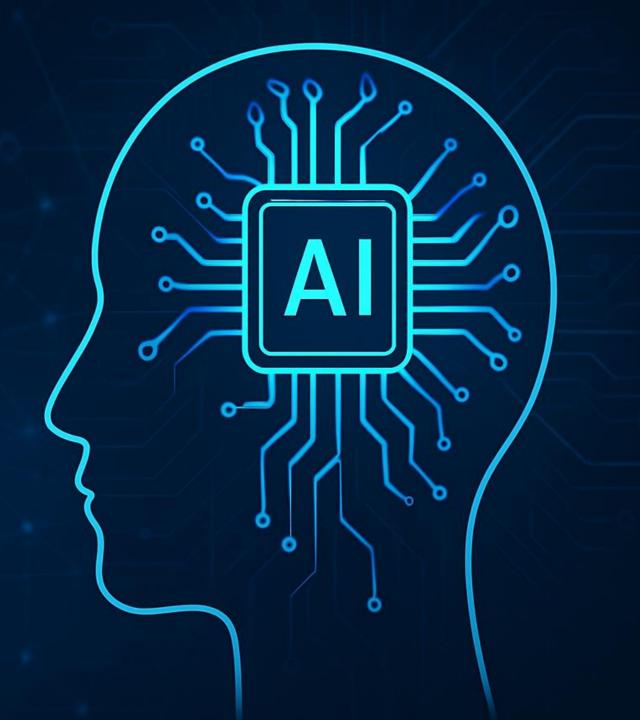
Current knowledge and availability of mach learning across the spectrum of trauma sc

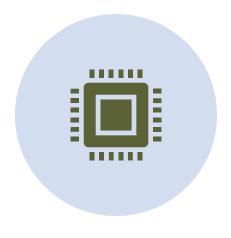
Tobias Gauss^a, Zane Perkins^b and Thorsten Tjardes^c

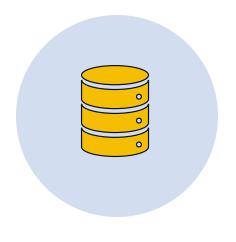
Study dataset (n derivation set)	Objective	Algorithm type	Features	Output
Maurer et al. [25 ^a] ACS-TQIP database (n= 934053)	Smartphone-based mortality predictor tool	Optimal Classification Trees	sBP, HR, SpO ₂ , RR, GCS, temperature, AIS, mechanism, comorbidities, demographics,	Mortality, organ failure, complications
Lee <i>et al.</i> [26] Hospital data (n=2232)			Age, HR, RR, MAP, GCS, AIS, ED interventions	7-day mortality
Nederpelt <i>et al.</i> [27] ACS-TQIP database (n=29816)	In-field triage tool for need of critical resources	Deep Neural Network	Age, BMI, HR, SBP, RR, Temperature, GCS, AIS	Shock, massive transfusion (MT), major surgery
Follin et al. [28] Traumabase registry (n = 1160)	Need for critical resources	Decision tree	Age, HR, SpO ₂ , SBP, GCS, ISS, mechanism	Critical interventions and resources
Liv <i>et al</i> . [29] Trial data (n=79)	Predict the need for lifesaving interventions in	Logistic regression (LR), Multilayer perceptron (MLP)	GCS, HR complexity, HR variability, shock index, pulse pressure, dBP, sBP, HR, SpO ₂	Life-saving intervention
Perkins <i>et al.</i> [18 ^a] Hospital data (n=1091)	Prediction coagulopathy, critical interventions, massive transfusion	Bayes network	HR, SBP, temperature, hemothorax, FAST, GCS, lactate, pH, mechanism, fractures	Coagulopathy, critical intervention, LOS ICU
Moyer et al. [30] Traumabase registry (n=2159)	Need for emergency neurosurgery	Logistic regression, Cat Boost	Prehospital physiological parameters and interventions	Need for emergency neurosurgery



MOHY NOW?







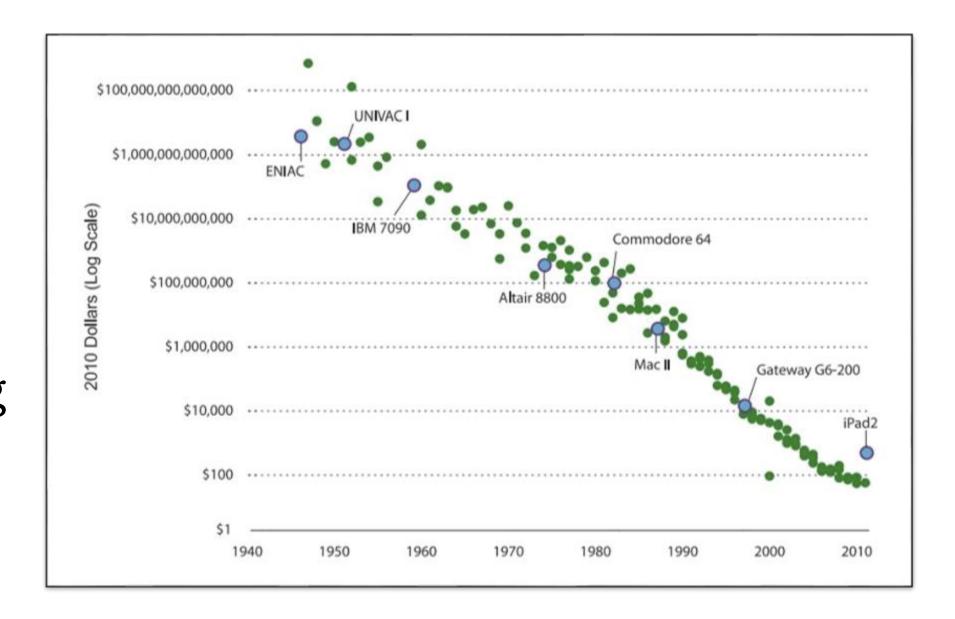


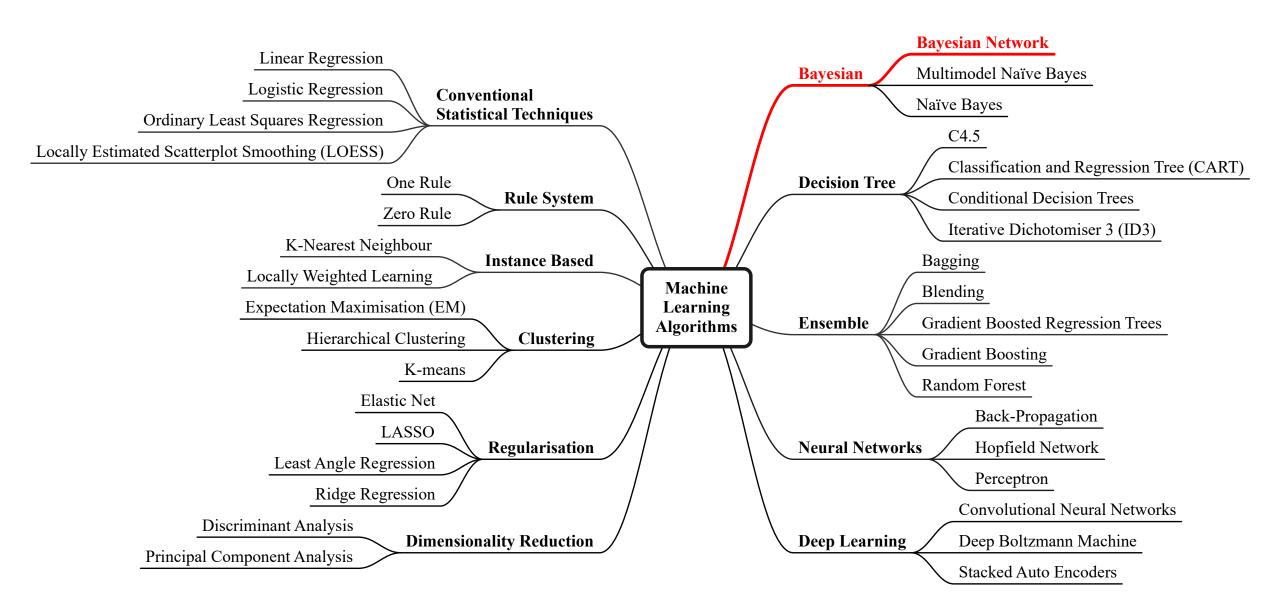
Compute

Data

Funding

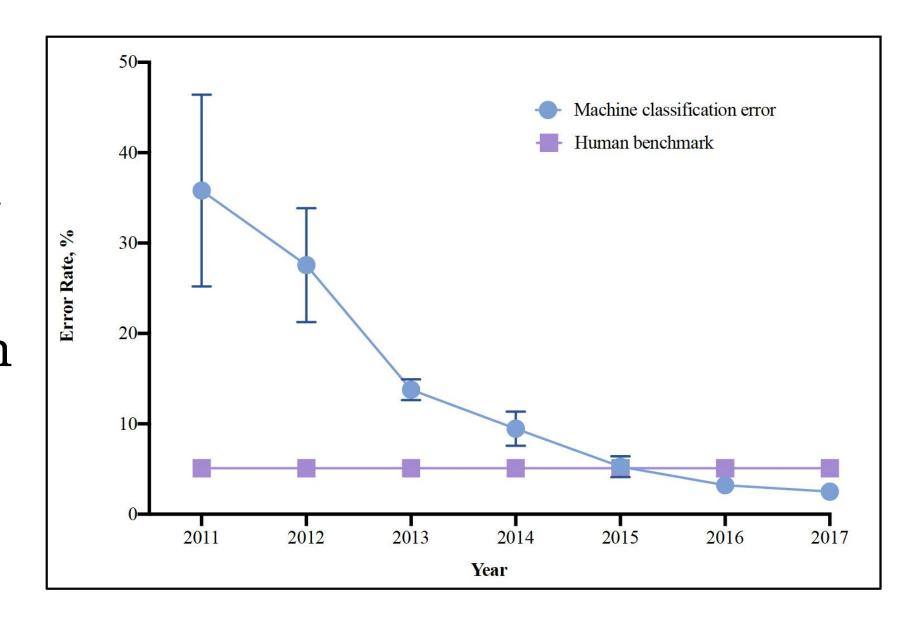
Cost of iPad 2 equivalent computing



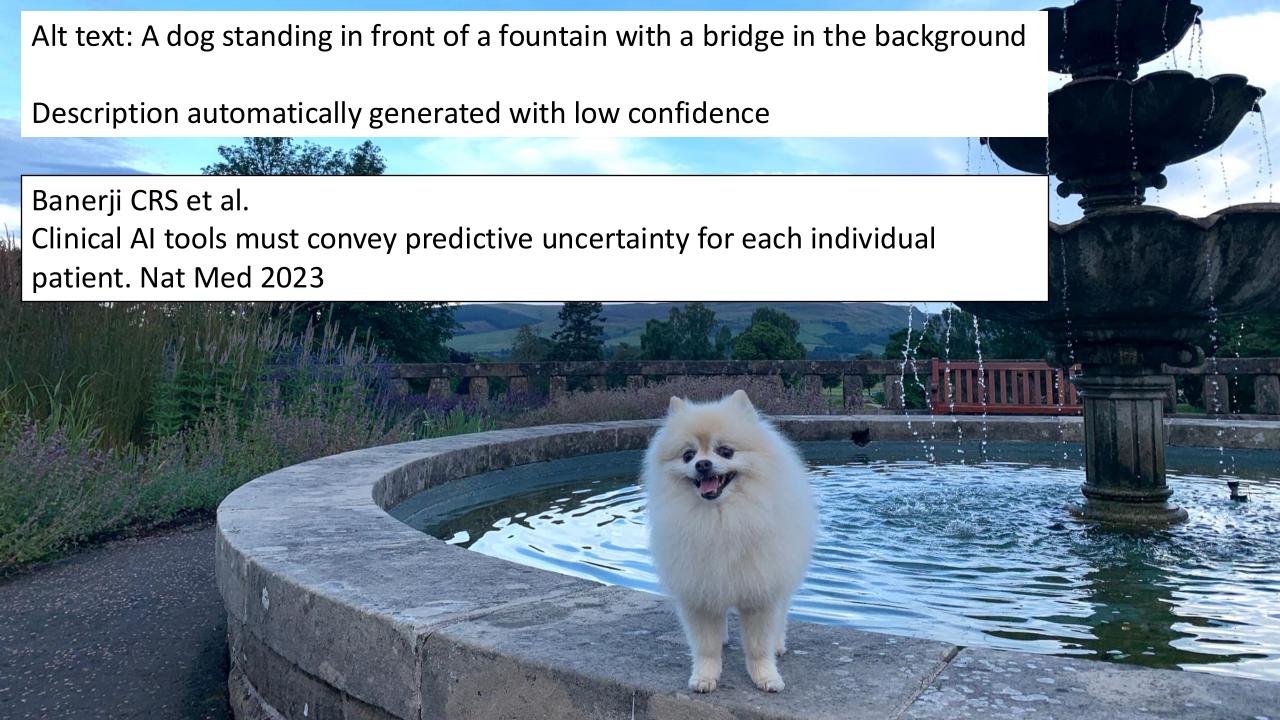


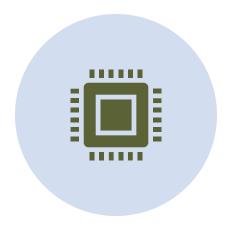
Error rate of image classification

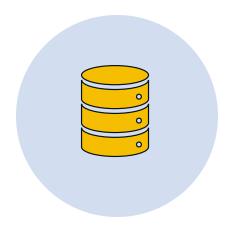
ImageNet challenge













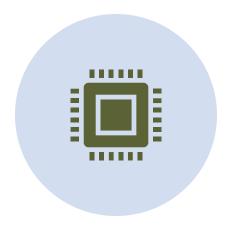
Compute

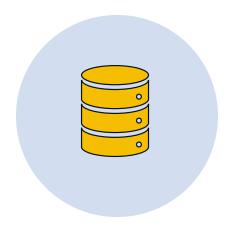
Data

Funding

"when you go from 10,000 training examples to 10 billion training examples, it all starts to work. Data trumps everything."

Google translate engineer





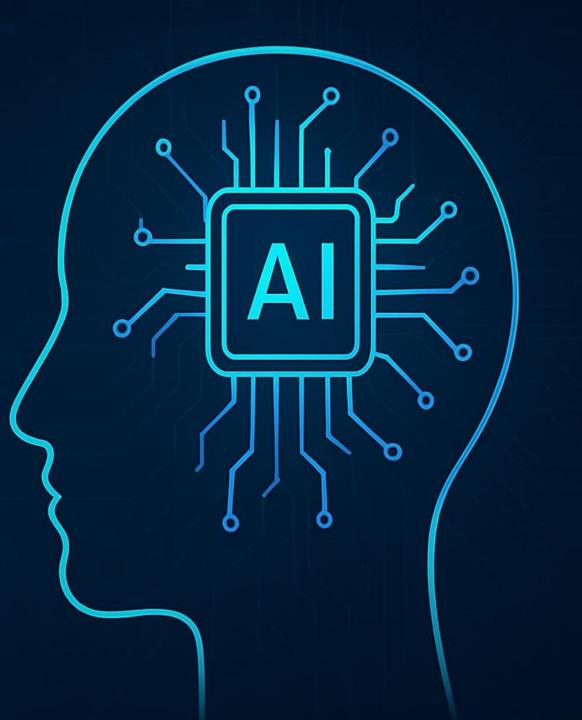


Compute

Data

Funding

CAUTION



bmjmedicine



¹Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, UK ²Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK ³Qatar University College of Medicine, Doha, Qatar ⁴Yale University, New Haven, Connecticut, USA Correspondence to: Professor Trisha Greenhalgh, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK; trish.greenhalgh@phc.ox.ac.uk

Cite this as: *BMJMED* 2025;4. doi:10.1136/bmjmed-2025-001394

Received: 4 February 2025 Accepted: 19 March 2025

How to read a paper involving artificial intelligence (AI)

Paul Dijkstra , ¹ Trisha Greenhalgh , ² Yosra Magdi Mekki, ³ Jessica Morley ⁴

ABSTRACT

This paper guides readers through the critical appraisal of a paper that includes the use of artificial intelligence (AI) in clinical settings for healthcare delivery. A brief introduction to the different types of AI used in healthcare is given, along with some ethical principles to guide the introduction of AI systems into healthcare. Existing publication guidelines for AI studies are highlighted. Ten preliminary questions to ask about a paper describing an AI based decision support algorithm are suggested.

Introduction

Increasingly, clinical research papers describe the use of artificial intelligence (AI), and much has been written about the potential of AI to revolutionise healthcare. Many examples exist of AI being rolled out into healthcare delivery (examples in boxes 1 and 2, and recent systematic reviews 3), from the use of a machine learning algorithm to distinguish between a cough caused by a pulmonary tuberculosis and non-tuberculosis respiratory condition (box 1) to using generative AI to impresse designations.

critical appraisal for papers that do not include the study of AI in clinical care.⁵

What is artificial intelligence?

AI is a transformative technology capable of completing tasks that typically require human intelligence. AI is also an interdisciplinary area of research that spans, among other disciplines, computer science, psychology, linguistics, and philosophy. AI in healthcare is divided into two broad categories:

- ▶ Artificial narrow intelligence refers to machine learning algorithms that can recognise patterns in large datasets. Artificial narrow intelligence is useful for solving text, voice, or image based classification and clustering problems, excelling at a precisely defined, single task (eg, playing chess).
- ▶ Artificial general intelligence refers to AI applications that can reason, argue, memorise, and solve problems. Artificial general intelligence is sometimes referred to as human level AI, because it is considered to display a cognitive capacity which approximates to that of a human being.

Whereas artificial narrow intelligence is already

10.1136/bmjmed-2025-001394 on 14 April 2025. Protected by copyright, including : ₫ related Downloaded q

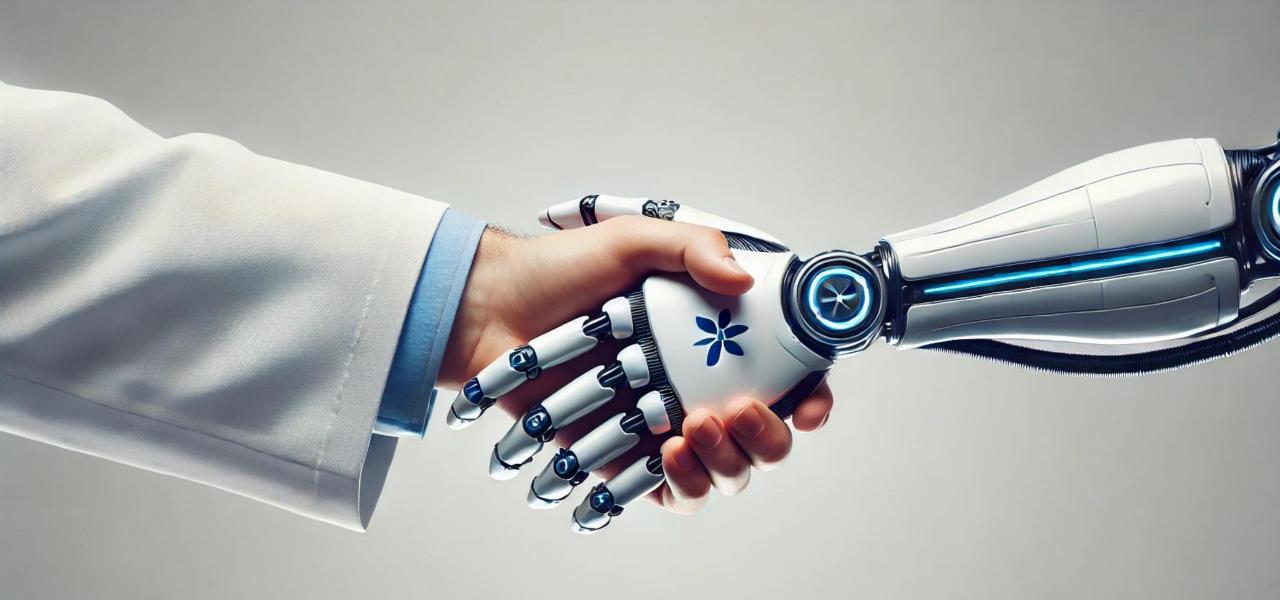
BMJ Medicine:

first published as

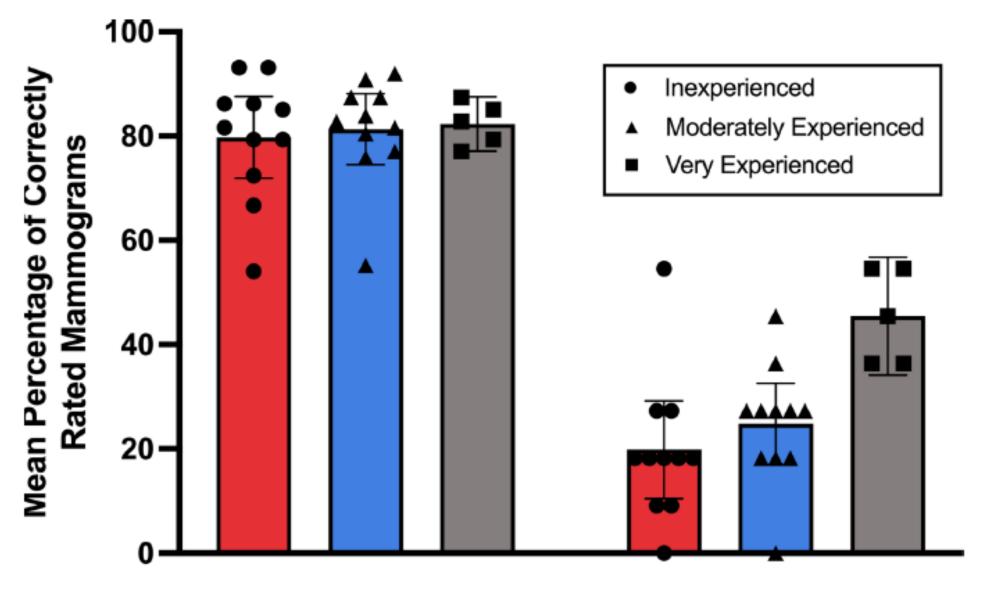


136/hmimed-2025-001394 on 14 April 2025 Downloade





Human Machine Pairing

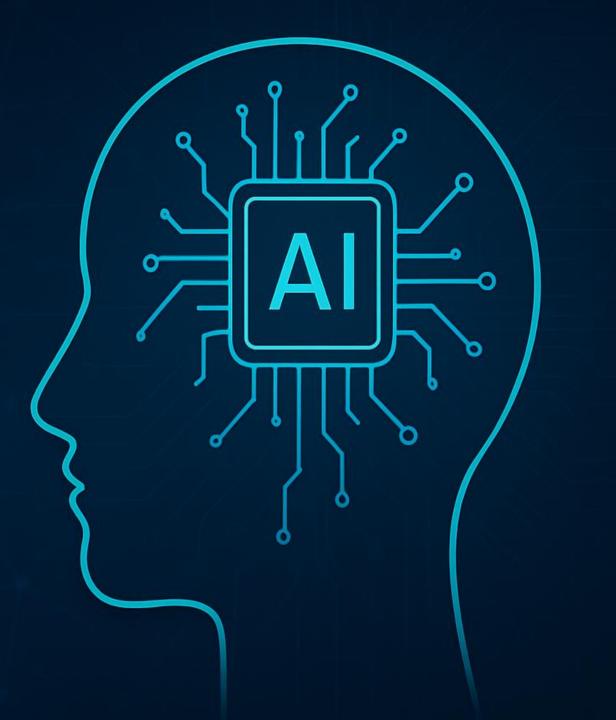


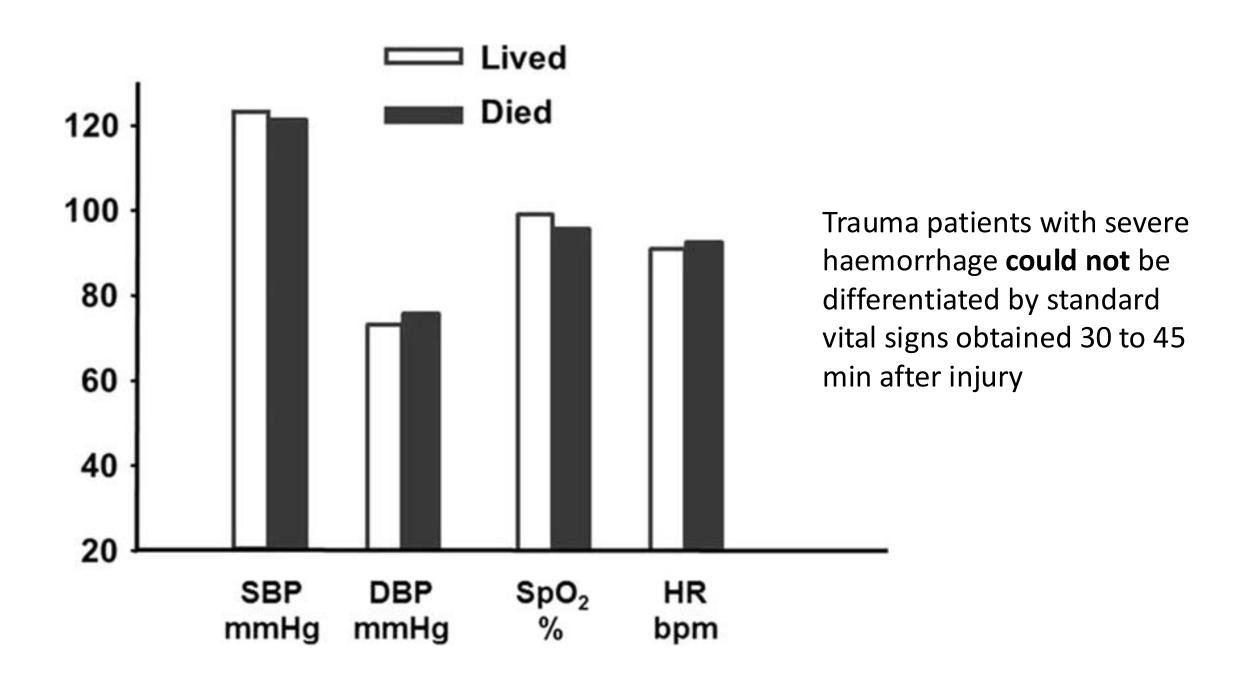
Al suggests correct BI-RADS Category Al suggests incorrect BI-RADS Category

Correctness of AI Suggestions

Dratsch, Automation Bias in Mammography *Radiology*, 2023

Blood prediction







Contents lists available at ScienceDirect

Injury





Clinical gestalt and the prediction of massive transfusion after trauma*



Matthew J. Pommerening a,b, Michael D. Goodman c, John B. Holcomb a,b, Charles E. Wade a,b, Erin E. Fox b,d, Deborah J. del Junco b,d, Karen J. Brasel e, Eileen M. Bulger f, Mitch J. Cohen g, Louis H. Alarcon h, Martin A. Schreiber j, John G. Myers j, Herb A. Phelan k, Peter Muskat hon behalf of the PROMMTT Study Group Bryan A. Cotton a,b,* MPH on behalf of the PROMMTT Study Group 1

"Is the patient likely to be massively transfused?"

All patients received >1 unit of blood 966 patients 23% had a MT



Injury

journal homepage: www.elsevier.com/locate/injury

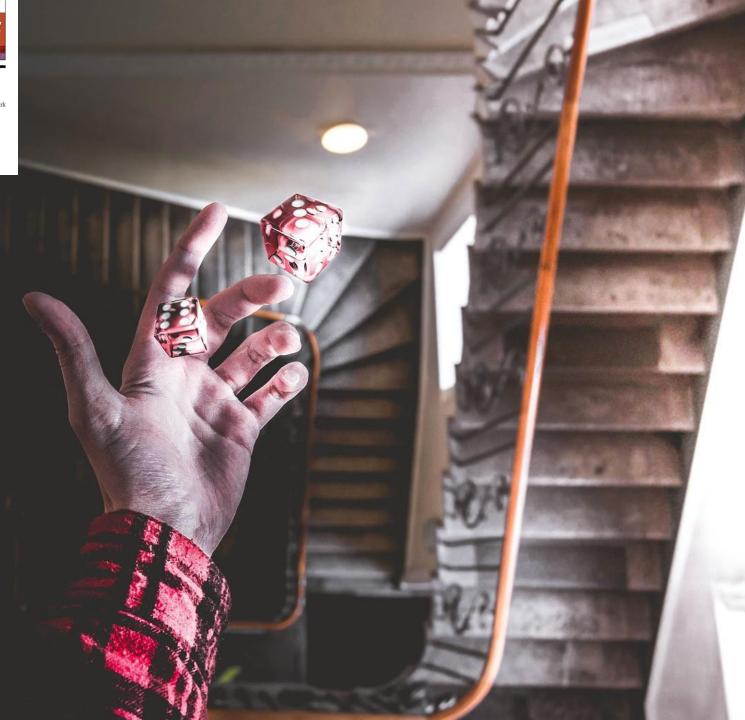
Clinical gestalt and the prediction of massive transfusion after trauma*

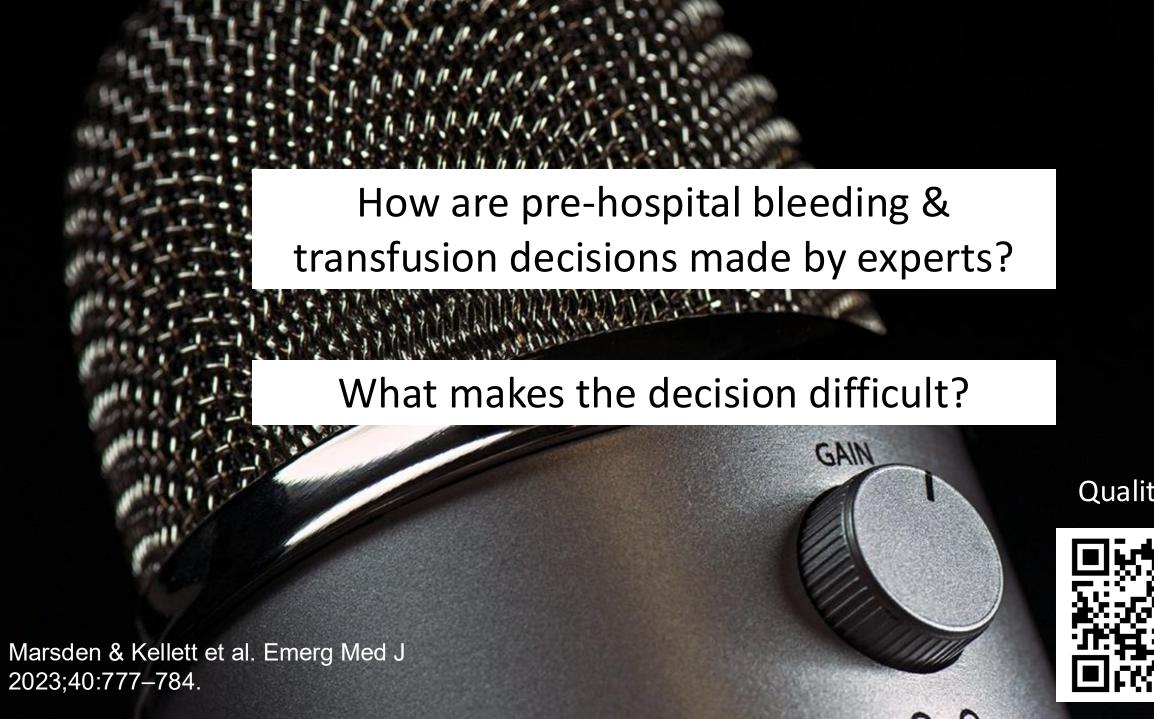
Matthew J. Pommerening ^{a,b}, Michael D. Goodman ^c, John B. Holcomb ^{a,b}, Charles E. Wade ^{a,b}, Erin E. Fox ^{b,d}, Deborah J. del Junco ^{b,d}, Karen J. Brasel ^e, Eileen M. Bulger ^f, Mitch J. Cohen ^g, Louis H. Alarcon ^h, Martin A. Schreiber ⁱ, John G. Myers ^j, Herb A. Phelan k, Peter Muskat C, Mohammad Rahbar b.l.

Bryan A. Cotton a,b,* MPH on behalf of the PROMMTT Study Group¹



Sensitivity 66% Specificity 64% PPV 35%





Qualitative



Trauma Surgery & Acute Care Open

Identification of major hemorrhage in trauma patients in the prehospital setting: diagnostic accuracy and impact on outcome

```
Jared M Wohlgemut , , , Erhan Pisirir , , Rebecca S Stoner , , , , Evangelia Kyrimi, Michael Christian, Thomas Hurst, William Marsh , , , Zane B Perkins , , , , Nigel R M Tai , ,
```

Quantitative



Open access Original research

Trauma Surgery & Acute Care Open Identification of major hemorrhage in trauma patients in the prehospital setting: diagnostic accuracy and impact on outcome

Jared M Wohlgemut , ^{1,2} Erhan Pisirir , ³ Rebecca S Stoner , ^{1,2} Evangelia Kyrimi, ³ Michael Christian, ⁴ Thomas Hurst, ⁴ William Marsh , ³ Zane B Perkins , ^{1,2} Nigel R M Tai^{1,2}

Accuracy of MHPA Sens 70%, Spec 94%

Mortality associated with missed Major Haemorrhage OR 3.3





Traditional Modelling Approaches

In 2020 there were >40 models for prediction of haemorrhage after injury

Many share input variables e.g. SBP, bloods and radiology

	АВС	TASH	Mina	Shock Index	Code Red
Publication Year	2009	2006	2013	2011	2016
Intended Location	IH	IH	IH	PH	PH
Internal validation (n)	596	6044	13961	n/a	126
External validation (n)*	5147	5147	1245	535	53
Impact study performed	retrospective	No	No	No	No
AUROC in external validation	0.76	0.89	~0.70	0.77	unknown
Major haemorrhage definition	1	2	3	1	4



Current knowledge and availability of machine learning across the spectrum of trauma science

Tobias Gauss^a, Zane Perkins^b and Thorsten Tjardes^c



Prediction of massive transfusion is overrepresented.

Received vs required



Predicted interventions correspond to therapeutic actions that clinicians have performed in the derivation cohorts. In consequence, the models predict past physician behaviour

Jöurnal of Trauma and **Acute Care Surgery**

Articles & Issues ✓ Collections ✓

Multimedia 🗸

For Authors >

Accurately predicting the need for blood transfusion following injury can be challenging.



Machine learning can be used to develop accurate prediction



25 machine learning models were identified.

> 8 models have excellent predictive performance.

4 models have been externally validated in hospital and prehospital settings.





Machine learning models for blood transfusion prediction in trauma are diverse with promising performance.



Future research must aim to bridge the gap between development and clinical application.



Cite



Share



SYSTEMATIC REVIEW W/O META-ANALYSIS

Predicting blood transfusion following traumatic injury using machine learning models: A systematic review and narrative synthesis

Oakley, William MSc¹; Tandle, Sankalp MBChB²; Perkins, Zane PhD, FRCS²; Marsden, Max PhD, FRCS¹





Predicting blood transfusion following traumatic injury using machine learning models: A systematic review and narrative synthesis

Oakley, William MSc1; Tandle, Sankalp MBChB2; Perkins, Zane PhD, FRCS2; Marsden, Max PhD, FRCS1

25 machine learning models were identified.

8 models have excellent predictive performance.

4 models have been externally validated in hospital and prehospital settings.





- Models had between 3 and 23 input variables
- Only 2 models reported measures of calibration



Predictors



Vital Signs 24 models



Demographics 12 models



Photoplethysmography
Waveforms
5 models

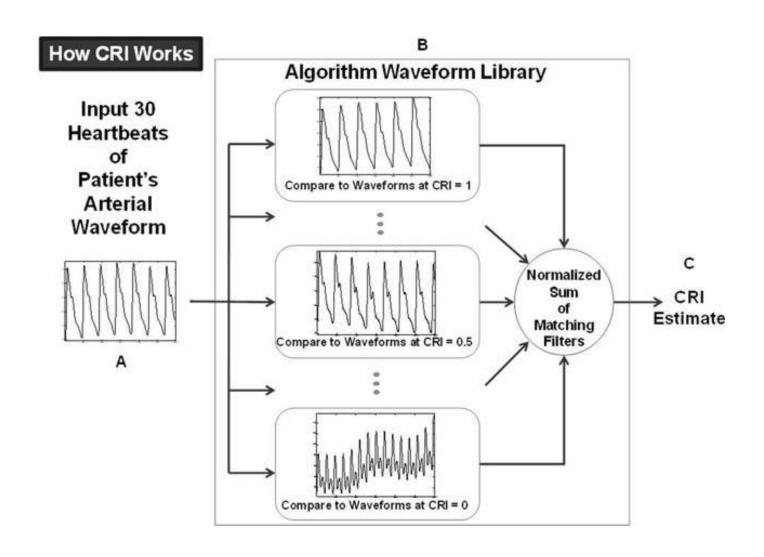


Laboratory Results 12 models



CT Interpretation 1 model

Feature Extraction



24 different outcomes

Blood Transfusion

Transfusion within 2 hours

Transfusion within 3 hours

Transfusion within 4 hours

Transfusion within 6 hours

Transfusion within 12 hours

Transfusion within 24 hours

Transfusion within 48 hours

Transfusion within 72 hours

Transfusion of universal donor Group O un-crossmatched PRBC

Prediction of extensive blood transfusion (>350ml)

Rapid transfusion (>5 units in 4 hours)

Massive Transfusion

10 units PRBC in 24 hours

4 units PRBC in 4 hours

6 units PRBC in 6 hours

Critical Administration Threshold (3 units PRBC in 60 minutes)

Resource Intensity greater than or equal to 4 5 units or greater PRBC within 5 hours of admission

10 units of greater PRBC within first 12 hours

> 10 units PRBC within 4 hours of admission>4 units in 1 hour within anticipated ongoing clinical need

Composite

Lifesaving Intervention

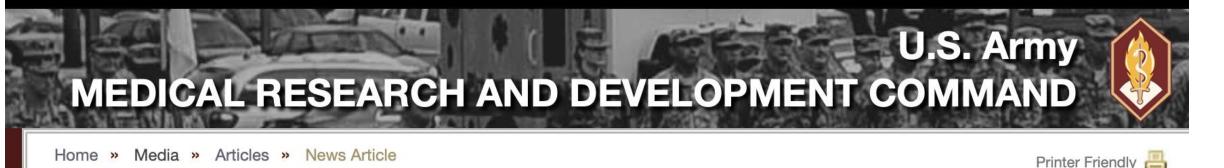
Haemorrhage requiring therapy or haemorrhage control intervention

Prediction of angioembolization, pelvic packing or massive transfusion

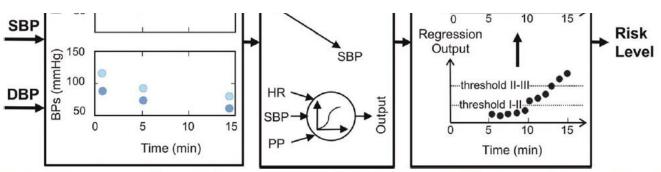
Prediction of shock



APPRAISE-HRI: AN ARTIFICIAL INTELLIGENCE ALGORITHM FOR TRIAGE OF HEMORRHAGE CASUALTIES



FDA Clears First Al Software for Hemorrhage Triage of Combat Casualties



AISE-HRI algorithm receives three vital signs as inputs (HR SRP and DRP) and provides a risk level as its or



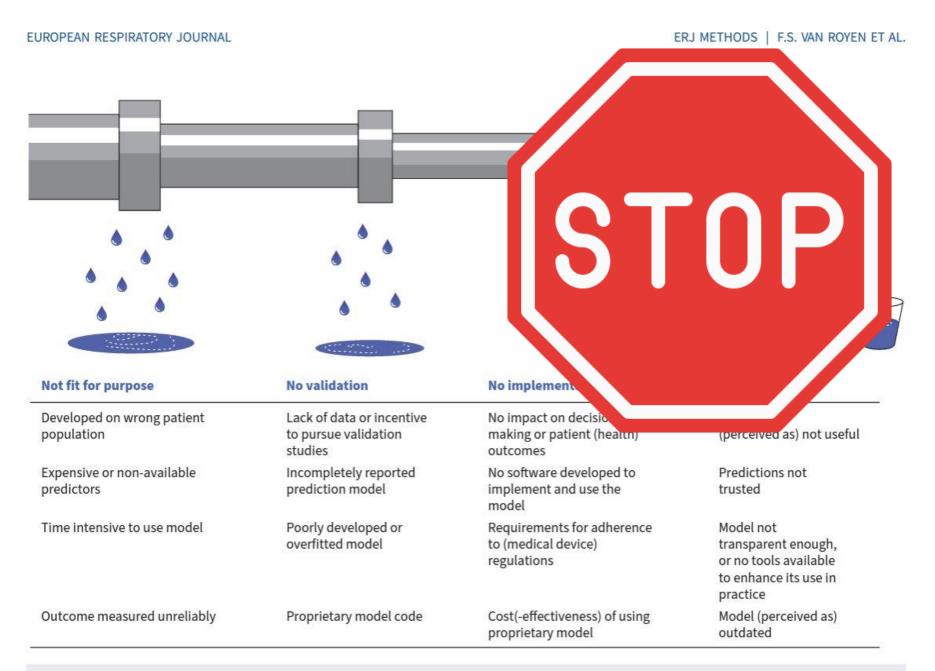
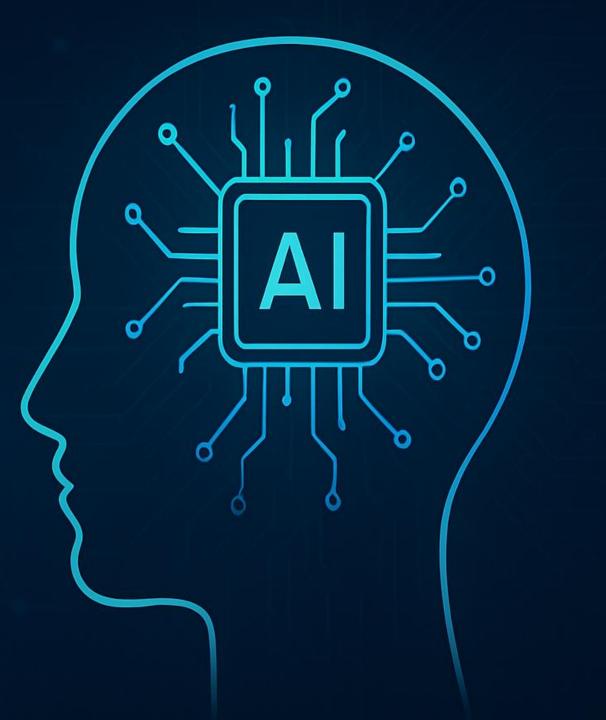
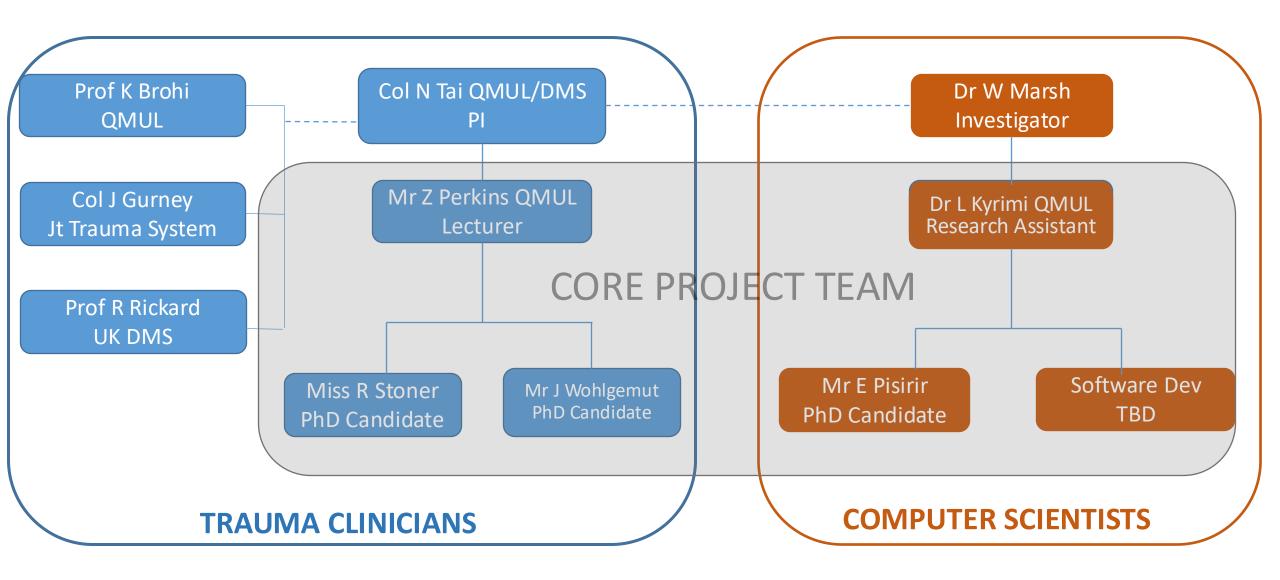


FIGURE 1 Leaky prognostic model adoption pipeline. Examples of reasons for failed prediction model adoption in clinical practice.

Bayesian Networks

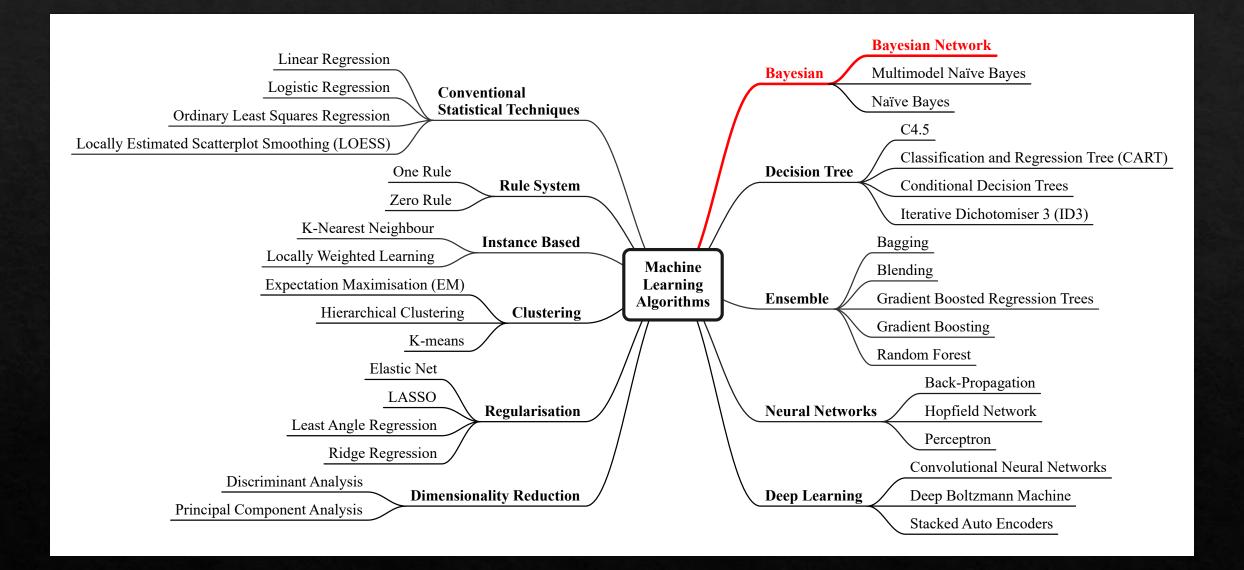




COMBAT-AID DEVELOPMENT TEAM

COMPUTER BATTLEFIELD ASSISTANCE IN TRAUMA CARE & INJURY DECISION SUPPORT

SUPPORTED BY US DOD W81XWH-19-2-0047 - Exploitation of Bayesian Networks for Clinical Decision Support on the Battlefield



Bayesian Network

Bayesian

Advantages of Bayesian Networks

Explicit modeling of uncertainty

Combine data with domain knowledge

Flexibility to update beliefs and incorporate new evidence

Ability to handle complex dependencies

Support inference in the presence of missing data

Support for probabilistic inference and reasoning under uncertainty

Support interventional and counterfactual reasoning

Explainable





Advantages of Bayesian Networks

Explicit modeling of uncertainty

Combine data with domain knowledge

Flexibility to update beliefs and incorporate new evidence

Ability to handle complex dependencies

Support inference in the presence of missing data

Support for probabilistic inference and reasoning under uncertainty

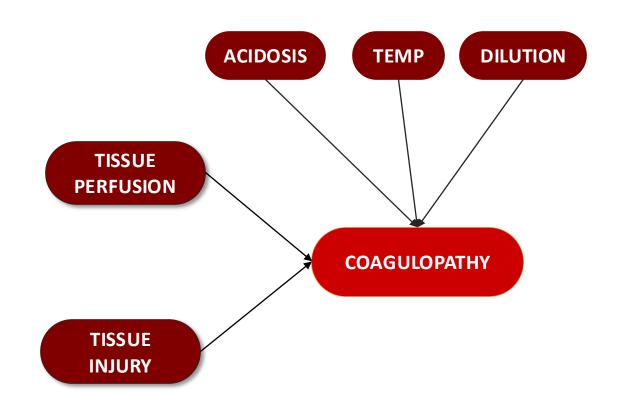
Support interventional and counterfactual reasoning

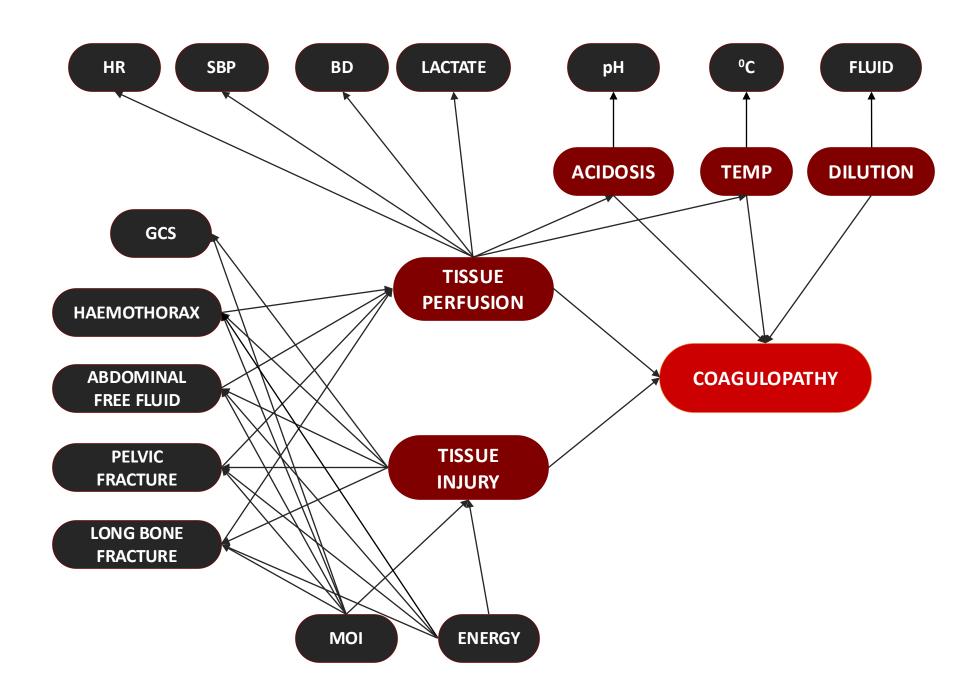
Explainable

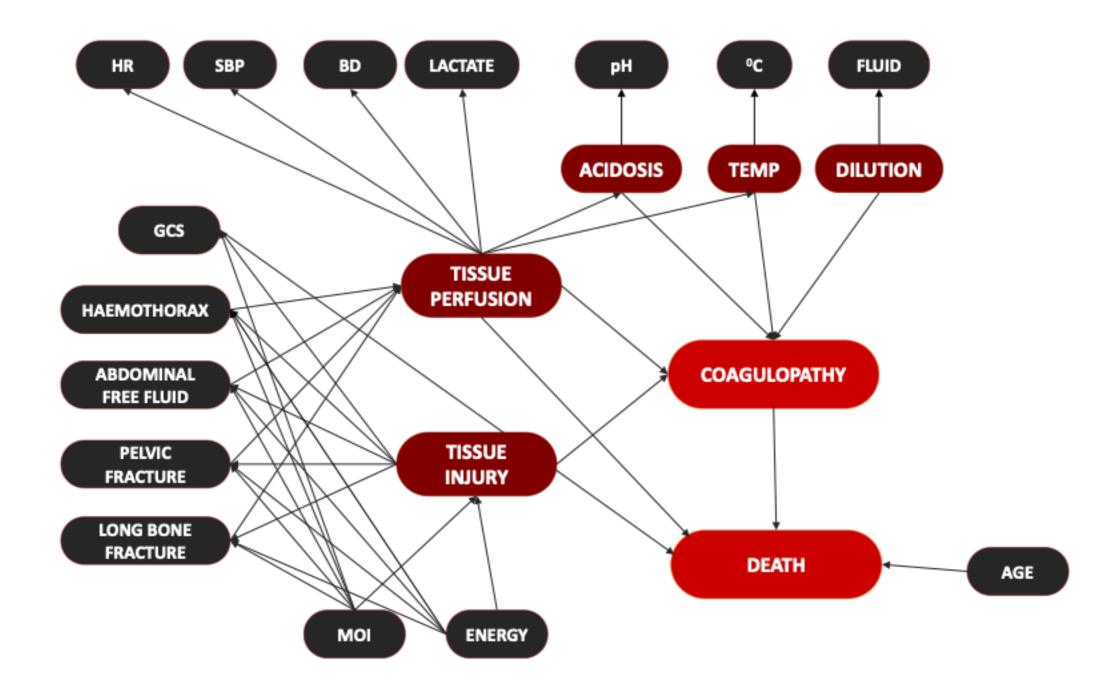




COAGULOPATHY







Journal of Biomedical Informatics 48 (2014) 28-37

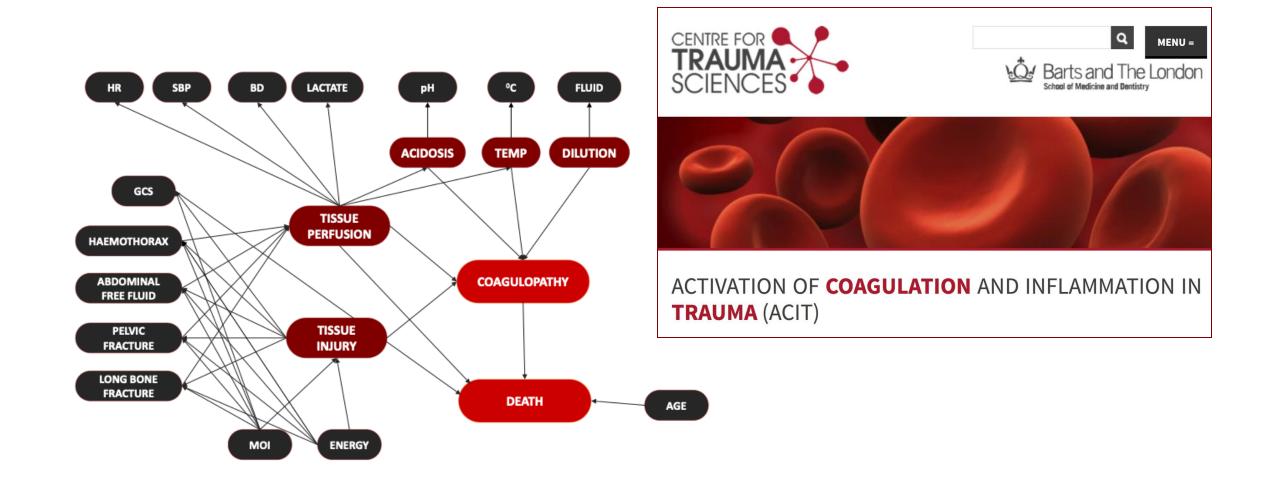
Not just data: A method for improving prediction with knowledge

Barbaros Yet ^{a,*}, Zane Perkins ^b, Norman Fenton ^a, Nigel Tai ^c, William Marsh ^a

Journal of Biomedical Informatics 52 (2014) 373–385

Combining data and meta-analysis to build Bayesian networks for clinical decision support

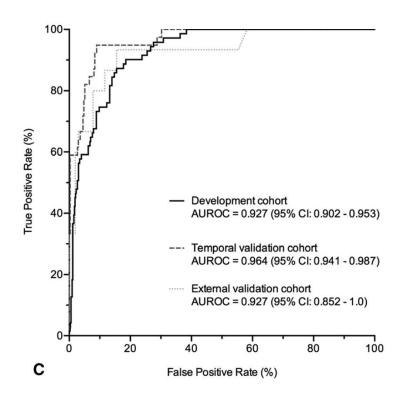
Barbaros Yet ^{a,*}, Zane B. Perkins ^b, Todd E. Rasmussen ^c, Nigel R.M. Tai ^d, D. William R. Marsh ^a

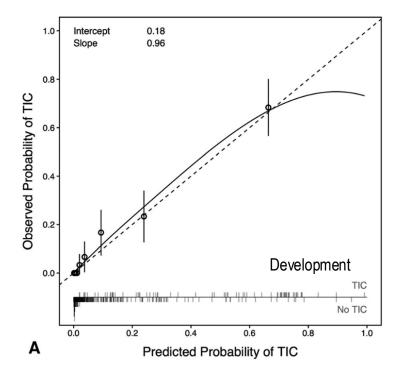


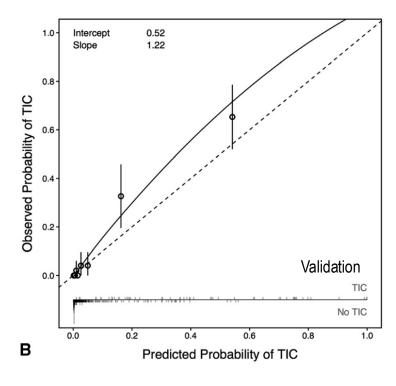
Early Identification of Trauma-induced Coagulopathy

Development and Validation of a Multivariable Risk Prediction Model

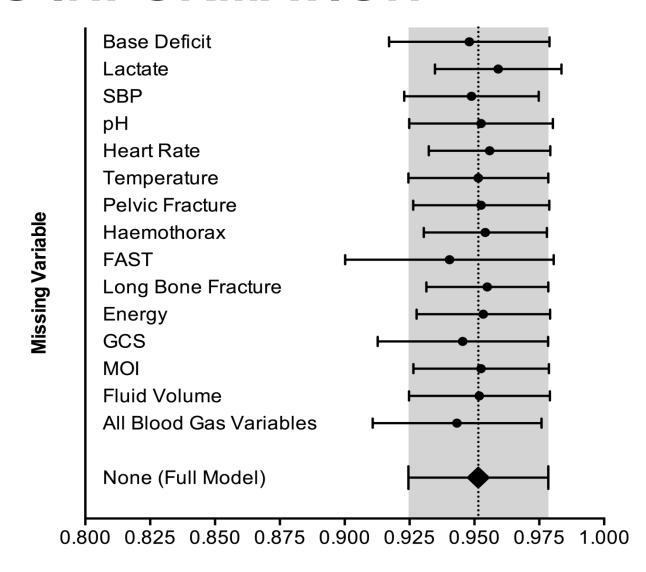
Zane B. Perkins, PhD,*⊠ Barbaros Yet, PhD,† Max Marsden, BSc,* Simon Glasgow, PhD,* William Marsh, PhD,† Ross Davenport, PhD,* Karim Brohi, MD,* and Nigel R. M. Tai, MD*‡







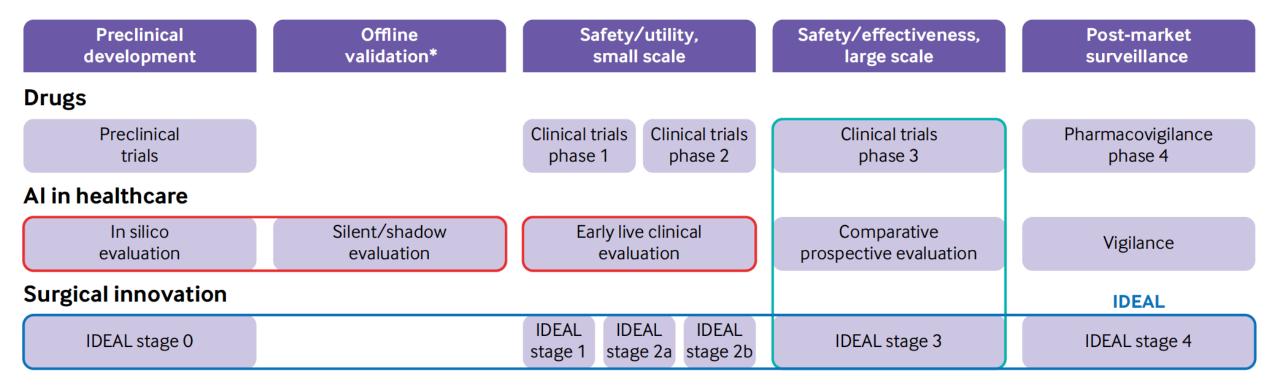
MISSING INFORMATION

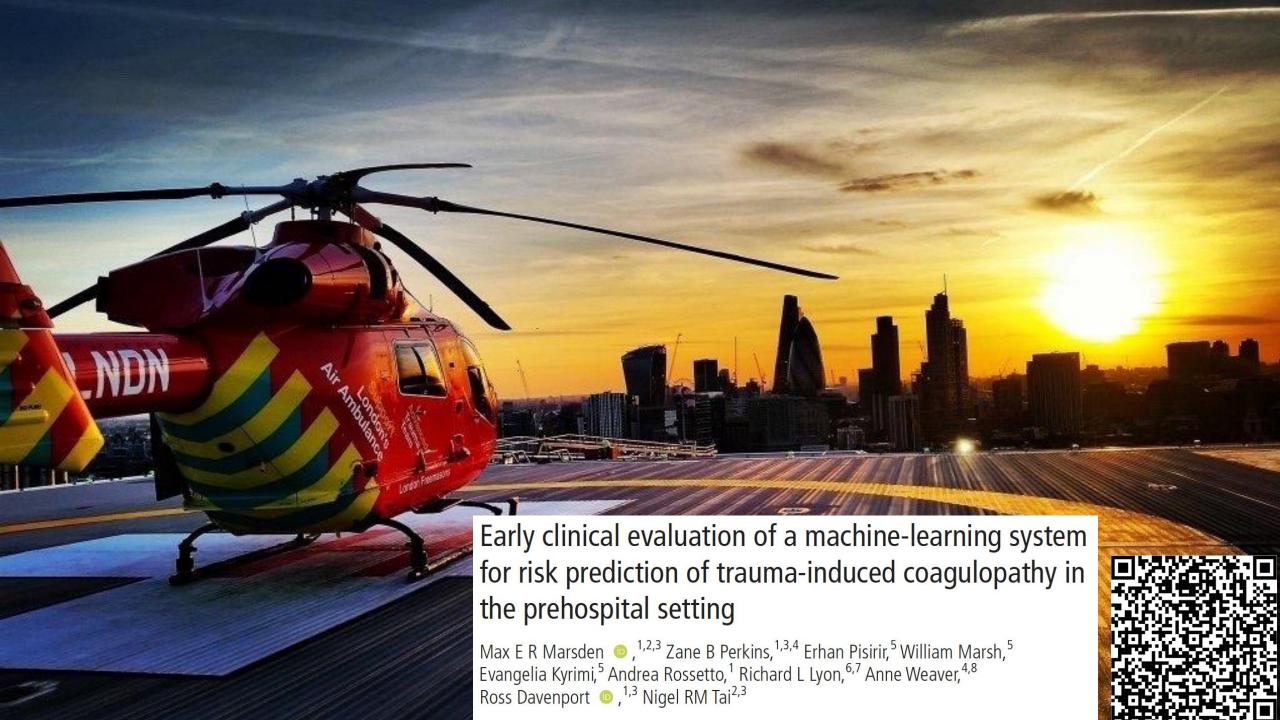


Area Under the ROC Curve

So what about impact?







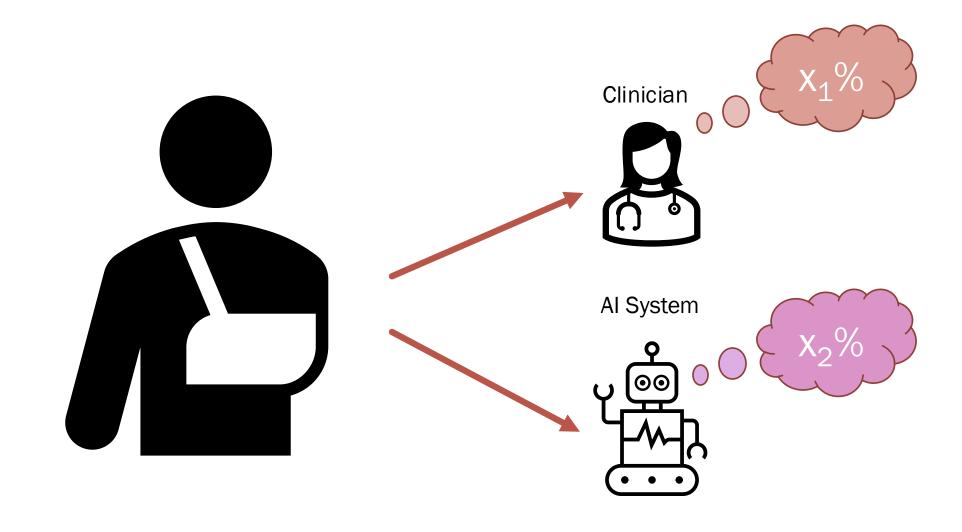


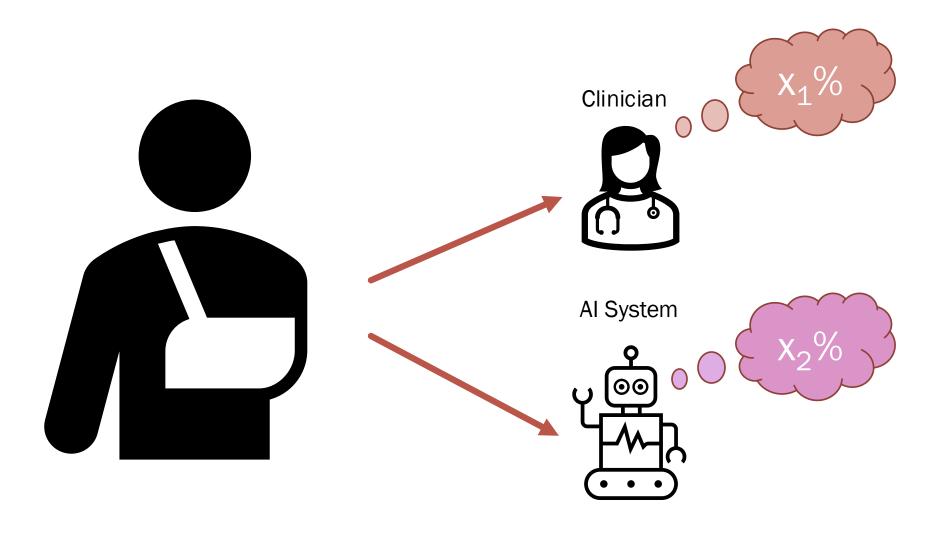
Method

Prospective multicentre impact study at 2 x Air Ambulances

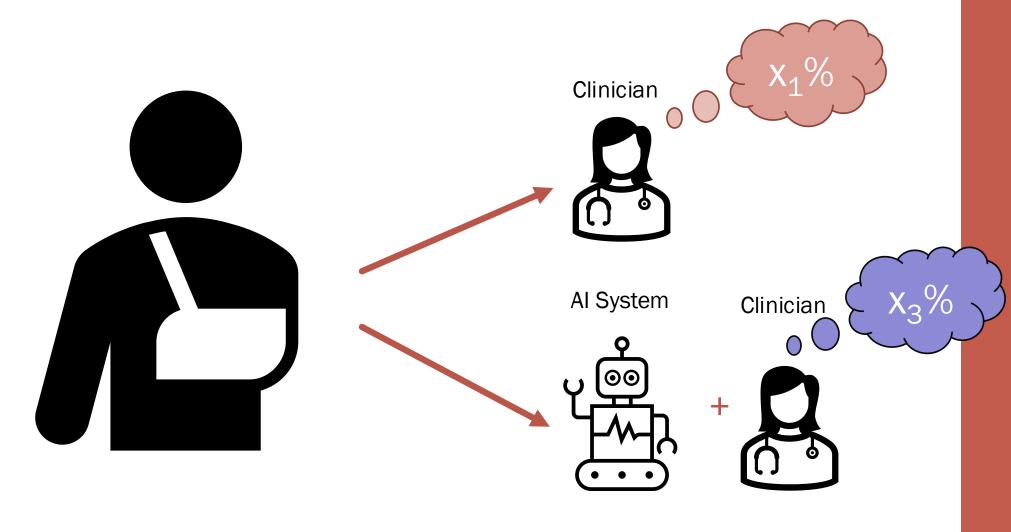
6 months 2018

Compared expert pre-hospital clinicians to an Al system



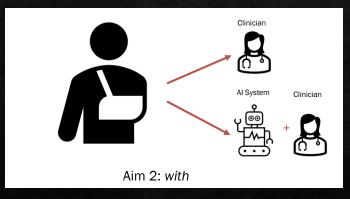


Aim 1: against

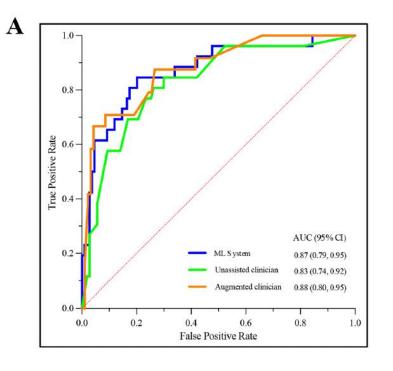


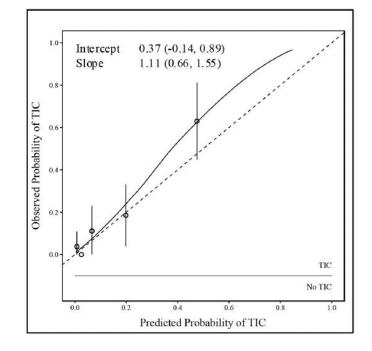
Aim 2: with

Risk prediction was better when clinicians were assisted with the AI system



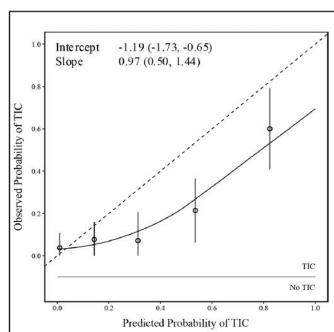
	Clinician	Al System & Clinician
Discrimination - AUROC	0.83 (0.74, 0.92)	0.88 (0.80, 0.95)
Calibration	-1.19 (-1.73, -0.65)	-0.79 (-1.38, -0.21)
Accuracy (Brier Skill)	0.00 (-0.41, 0.30)	0.27 (-0.05, 0.51)

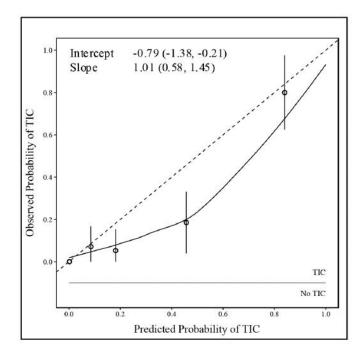




B

D





A: ML as good as Doctors (AUROC)

B: ML (Calibration) – tends to small under-triage

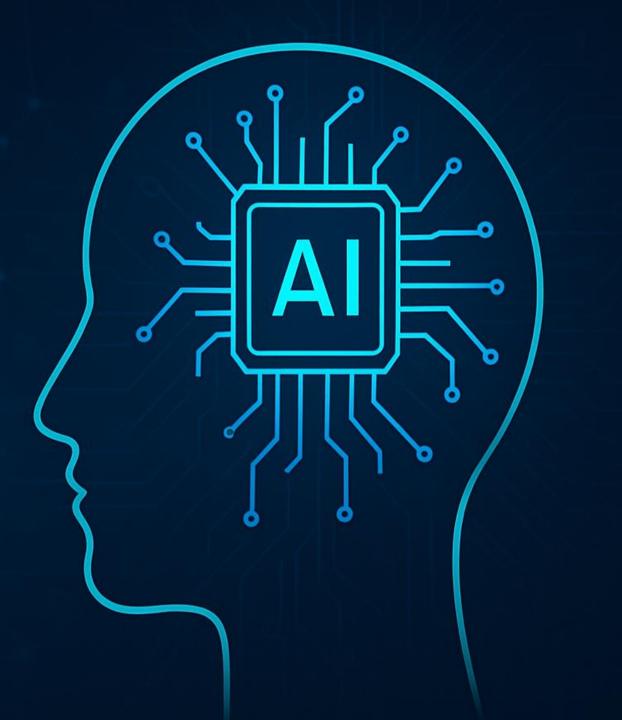
C: Docs (Calibration) – tend to moderate over-triage for "grey and sick"

D: ML+Docs (Calibration) = moderate over-triage for "grey" areas"

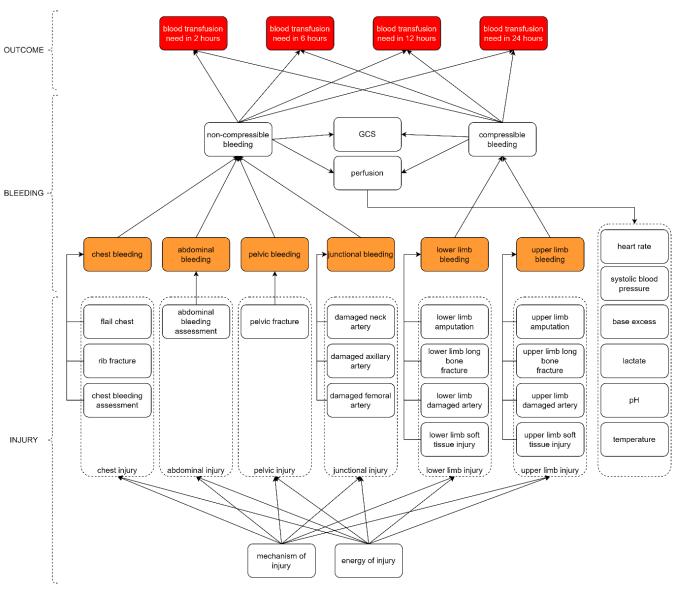


Human with Al produced the best predictions

Bleeding and Bayes



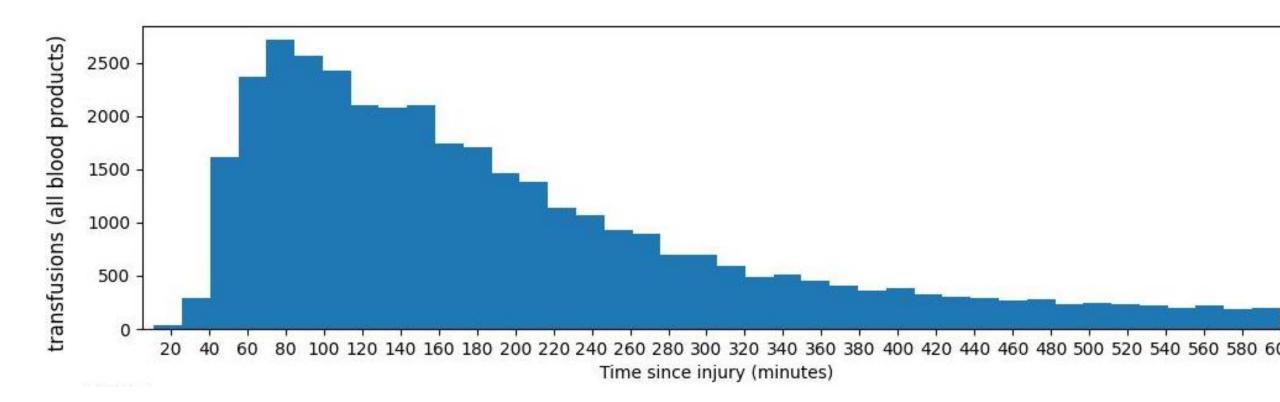
Predict the transfusion requirement at 2, 6, 12 and 24 hours
Using pre-hospital information



Training Dataset First 10 hours after injury

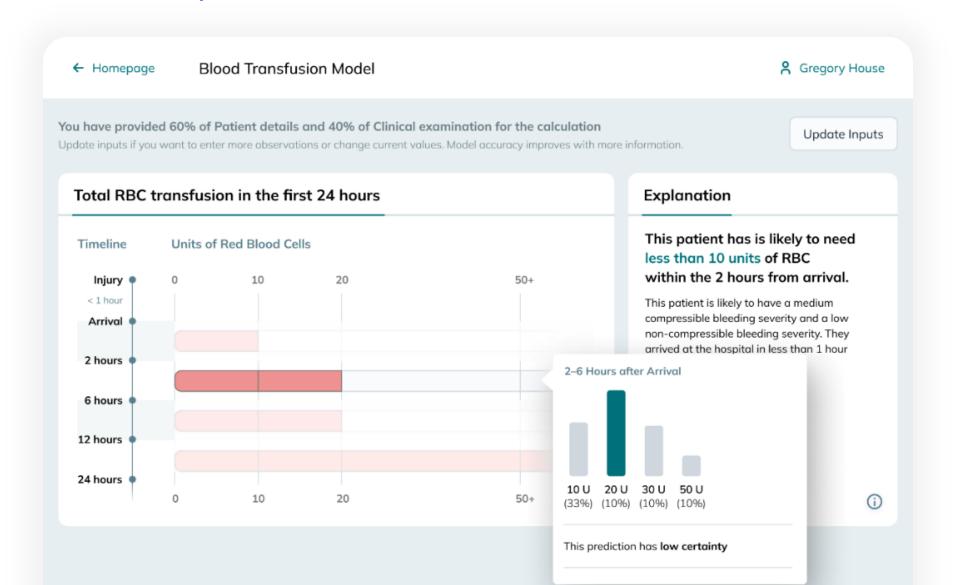
UK Joint Theatre Trauma Registry 2006 to 2014

36,294 units of blood products transfused to 2064 patients

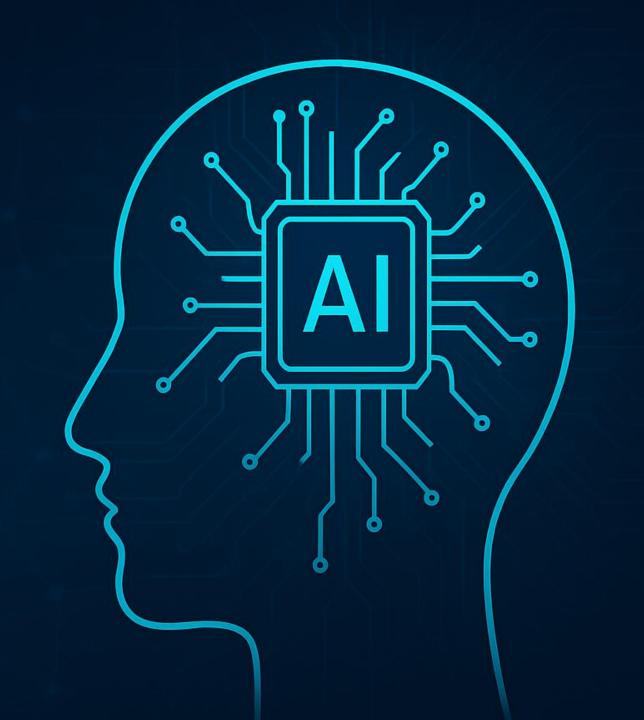


Our work: what is next?

From prediction algorithm to a decision support tool for clinicians



Future Trial

























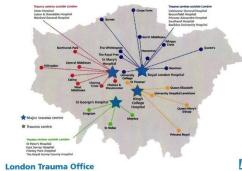






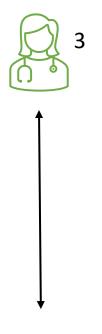
1. Sites: LAA, LAS, KCH, RLH, SMH, SGH

The London Trauma System









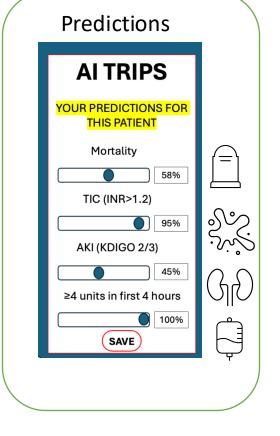


- 1. Sites: LAA, LAS, KCH, RLH, SMH, SGH
- 2. Trauma Patients retrieved by LAA and brought to MTC
- 3. Clinical decision maker in MTC ED









- 1. Sites: LAA, LAS, KCH, RLH, SMH, SGH
- 2. Trauma Patients retrieved by LAA and brought to MTC
- 3. Clinical decision maker in MTC ED
- 4. Capture of decision-maker predictions (Near Real Time)

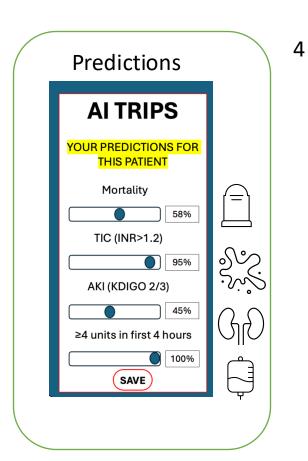


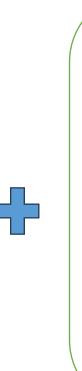


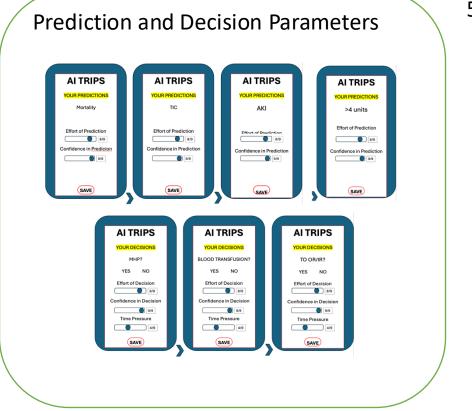








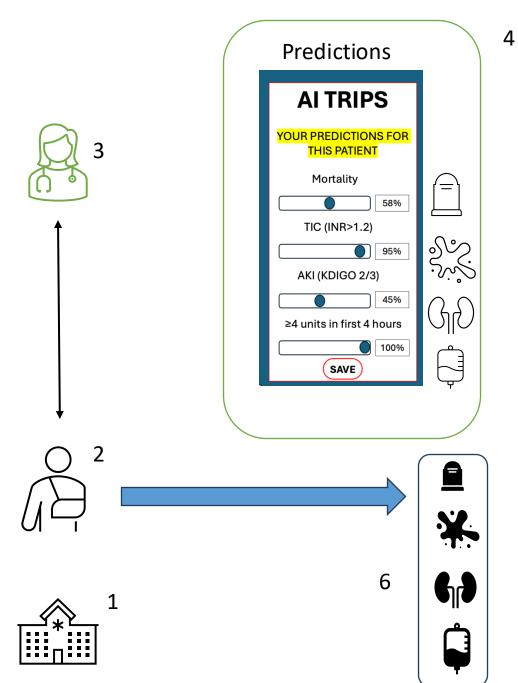


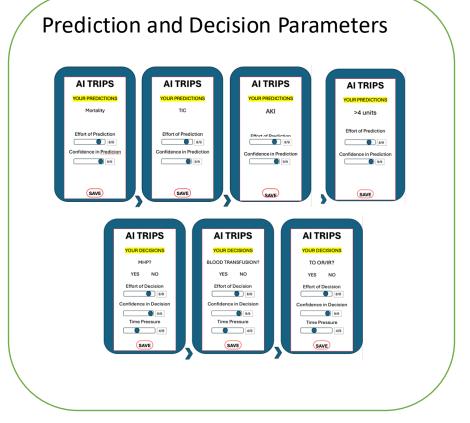


- 1. Sites: LAA, LAS, KCH, RLH, SMH, SGH
- 2. Trauma Patients retrieved by LAA and brought to MTC
- 3. Clinical decision maker in MTC ED
- 4. Capture of decision-maker predictions (Near Real Time)
- 5. Capture of prediction and decision parameters (Near Real Time)





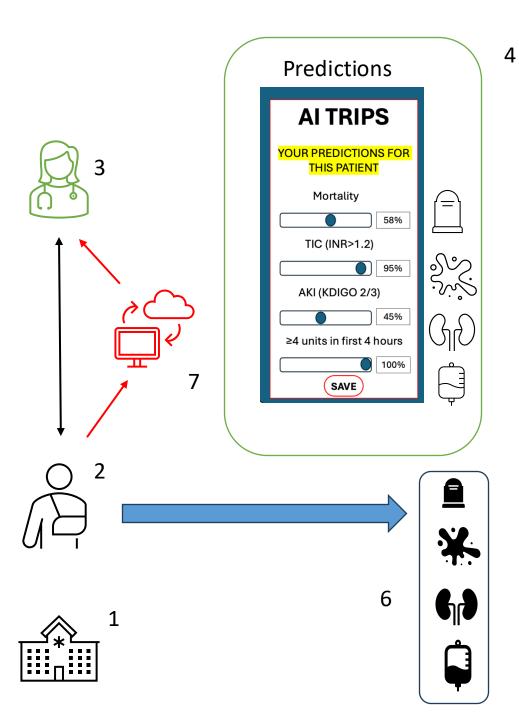


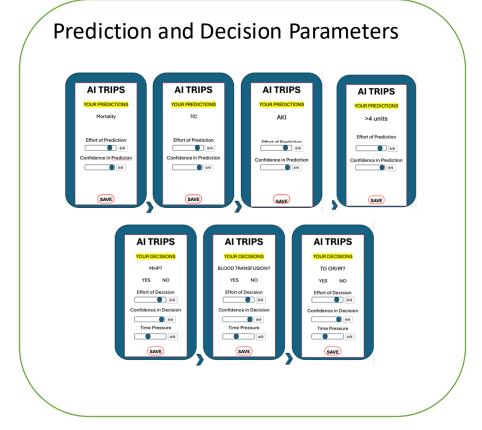


- 1. Sites: LAA, LAS, KCH, RLH, SMH, SGH
- 2. Trauma Patients retrieved by LAA and brought to MTC
- 3. Clinical decision maker in MTC ED
- 4. Capture of decision-maker predictions (Near Real Time)
- 5. Capture of prediction and decision parameters (Near Real Time)
 - Offline capture of patient outcome data from National Major Trauma Registry)



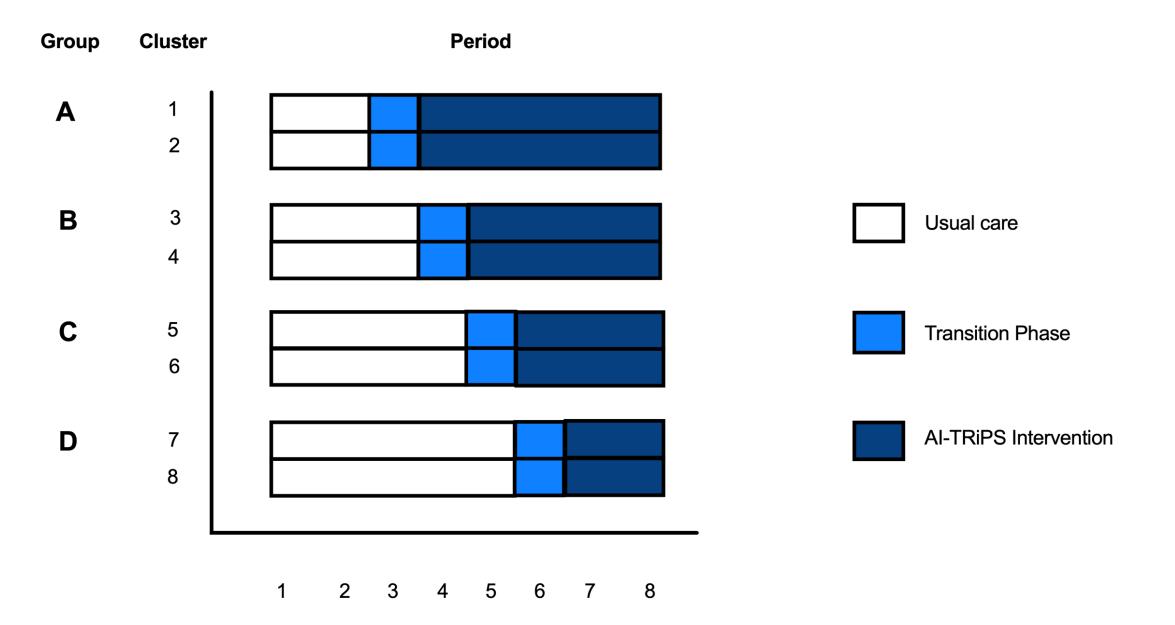






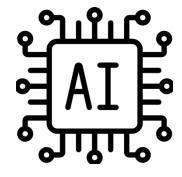
- 1. Sites: LAA, LAS, KCH, RLH, SMH, SGH
- 2. Trauma Patients retrieved by LAA and brought to MTC
- 3. Clinical decision maker in MTC ED
- 4. Capture of decision-maker predictions (Near Real Time)
- 5. Capture of prediction and decision parameters (Near Real Time)
- 5. Offline capture of patient outcome data from National Major Trauma Registry
- 7. Implement AI TRIPS Interface and continue 4/5/6













Advances due to computing power, data and funding

Al is increasingly prevalent

Need evidence to bridge the Al Chasm

Summary

THOR Solstrand Hotel, Norway 19-21st May 2025

Artificial Intelligence: Bleeding

Max Marsden

Honorary Clinical Lecturer, Centre for Trauma Sciences, QMUL, UK

Senior Lecturer in Surgical Informatics, Academic Department of Military Surgery and Trauma









