The Impact of a Bayesian Network for Pre-Hospital Decision-Support after Trauma

Max Marsden

Centre for Trauma Sciences
Blizard Institute
Queen Mary, University of London

2020

Submitted in partial fulfilment of the requirements of the
Degree of Doctor of Philosophy

Supervisors

Colonel Nigel Tai

Consultant Trauma and Vascular Surgeon
The Royal London Hospital
London, UK
Deputy Medical Director
Defence Medical Services
Litchfield, UK

Professor Karim Brohi

Professor of Trauma Sciences
Centre for Trauma Sciences
The Blizard Institute
Queen Mary, University of London
London, UK

Statement of Originality

I, Max Marsden, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below, and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Date: 30th October 2020

Collaboration:

Throughout the thesis I have been advised by Colonel Nigel Tai and Dr Zane Perkins. I am very grateful for the overarching support of Dr William Marsh and Professor Karim Brohi.

I have collaborated with the following people: Chapter 2: Data collection Dr Rahul Bagga and Dr Suzanne Kellett. Critical revision: Dr Katie Gillies and Dr Julia Williams. Chapter 4 and 5: Calibration curves: Dr Andrea Rossetto. Chapter 4: Decision analysis curve: Ms Mariana Raniere Neves. Chapter 6: Data organisation and initial analysis: Ms Amy Easthope. Chapter 7: The chapter is a systematic review based on a previous concept published in Dr Zane Perkins' PhD thesis. There is overlap between theses in the results section. The majority of the material is new and expands on previous work.

Associated Publications

Marsden MER, Mossadegh S, Marsh W, Tai N. Development of a major incident triage tool: the importance of evidence from implementation studies. J R Army Med Corps 2018. Online 26 Sept 2018

Marsden MER, Rossetto A, Duffield CAB, Woolley TGD, Buxton WP, Steynberg S, Bagga R, Tai, N. Prehospital tranexamic acid shortens the interval to administration by half in Major Trauma Networks: a service evaluation. Emerg Med J 2019;36:395-400.

Lendrum R, Perkins Z, Chana M, **Marsden MER**, Davenport R, Grier G, Sadek S, Davies G. Pre-hospital Resuscitative Endovascular Balloon Occlusion of the Aorta (REBOA) for exsanguinating pelvic haemorrhage. Resuscitation 2019;135:6-13.

Mclachlan S, Kyrimi E, Daley B, Kudakwashe D, **Marsden MER**, Finer S, Hitman G, Fenton N. Incorporating Clinical Decisions into Standardised Caremaps. 2020. Easychair Pre-print 2745

Perkins ZB, Yet B, **Marsden MER**, Marsh W, Tai N. Early identification of Trauma-Induced Coagulopathy: development and validation of a multivariable risk prediction mode. Ann Surg. 2020. Online 14 Jan 2020

Acknowledgements

So many people have supported me so I could complete this work. None more so than my Wife and Daughter. Words cannot express the love and gratitude I feel towards Tamsin and Sylvia. Without their unending support, I could never have written this thesis. I will forever be grateful for their patience with a deadline that was always just beyond touching distance.

I have been extremely lucky to work for an inspirational group of academics and clinicians. Nigel Tai has been a superb supervisor. His vision, advice and encouragement were invaluable to me. Equally, Zane Perkins and Karim Brohi have taught me so much about writing and research. I am very grateful to them all.

For technical assistance, I am grateful for the time and patience of Evangelia Kyrimi, Mariana Raniere Neves, Suzanne Kellet, Katie Gillies, Julia Williams, and Andrea Rossetto. I am indebted to the teams at the Air Ambulances both Kent, Surrey and Sussex through Richard Lyon and London, through Gareth Grier and David Lockey. It was a privilege to be welcomed into their institutions. I am very lucky to have received the time, funding and support from the British Army.

Finally, thanks to my parents, and parents-in-law, who got me to a place where I could submit a PhD thesis. I will always be grateful for the foundations they laid. Uncle-in-law Stuart Prebble, horribly overqualified, kindly read every word of this thesis and offered magnificent writing advice – thank you.

For everyone's hard work, help and sacrifice, my deepest hope is the findings in this thesis contribute to improving a patients' care in the future.

Thank you all for your help.

Abstract

In patients with major traumatic injuries, early intervention can be lifesaving. However, identifying high-risk patients can be difficult, and judgement errors may compromise optimal care. Prediction models can be used to augment clinical judgment. The aim of the thesis was to assess whether Bayesian Networks (BN), which are causal probabilistic models able to fuse knowledge and data, can augment clinical judgment in pre-hospital trauma care. The thesis focuses on decisions relating to Trauma Induced Coagulopathy (TIC), a difficult to diagnose condition that is central to resuscitation decisions, and haemorrhage.

To assess decision making, thematic analysis of interviews with expert clinicians was performed. The interviews revealed that situational recognition is key to effective decision making, and uncertain information results in difficult decisions. An example of uncertainty was evident in clinician's decisions to activate the major haemorrhage protocol. In such decisions, the predictive performance of an existing BN model ("ED TIC BN") was not significantly different to native clinical judgement (at equal specificity, sensitivity = clinicians: 86% vs BN: 96%, $p = 0.13$).

The ED TIC BN risk prediction model was updated for pre-hospital use to enable earlier decision support, and the *PH* TIC BN performed better than any other TIC model. The impact of the pre-hospital model on decision-making was analysed in a prospective, multicentre study. The discriminative performance of the BN was not significantly different from clinical assessment. The performance of clinical assessment did not improve after decision support information. However, there was evidence of the positive influence of the PH TIC BN on clinical assessment and increased ease of decision making.

This thesis advances the understanding of impact analysis in pre-hospital decision support after injury. The knowledge generated will be used to design subsequent trials to assess the impact of decision support on patient and health system outcomes.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAKSS	Air Ambulance Kent, Surrey & Sussex
ABC	Assessment of Blood Consumption
ACIT	Activation of Coagulation and Inflammation in Trauma
AI	Artificial Intelligence
AIS	Abbreviated Injury Scale
ANN	Artificial Neural Network
APPT	Activated Partial Thromboplastin Time
ATC	Acute Traumatic Coagulopathy
ATLS	Advanced Trauma Life Support
AUROC	Area Under the Receiver Operating Characteristic curve
BBV	Blood Borne Virus
BN	Bayesian Network
BD	Base Deficit
CA5	Clot Amplitude at 5 minutes
CDST	Clinical Decision Support Tool
CI	Confidence Interval
COAST	Coagulopathy of Severe Trauma
CRF	Case Report Form
CRI	Compensatory Reserve Index
DAG	Directed Acyclic Graph
DCA	Decision Curve Analysis
DCR	Damage Control Resuscitation
DMS	Defence Medical Services
ED	Emergency Department
FAST	Focus Assessment with Sonography for Trauma
FFP	Fresh Frozen Plasma
GCS	Glasgow coma scale
HB	Heuristics and Biases
HEMS	Helicopter Emergency Medical Services
HR	Heart Rate
INR	International Normalised Ratio
IQR	Inter-Quartile Range
ISS	Injury Severity Score
JTTR	Joint Theatre Trauma Registry
LAA	London's Air Ambulance
MERT	Medical Emergency Response Team
MHP	Major Haemorrhage Protocol
ML	Machine Learning
MODS	Multiple Organ Dysfunction Syndrome
MOI	Mechanism of Injury
MT	Massive Transfusion
MTC	Major Trauma Centre
NDM	Naturalistic Decision Making
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NISS	New Injury Severity Score
PACT	Prediction of Acute Coagulopathy of Trauma
PH	Pre-hospital
PT	Prothrombin Time
PRBC	Packed Red Blood Cells
RCT	Randomised Controlled Trial
RPD	Recognition Primed Decision
SBP	Systolic Blood Pressure
SI	Shock Index
SOP	Standard Operating Procedure
TARN	Trauma Audit Research Network
TBI	Traumatic Brain Injury
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
TIC	Trauma Induced Coagulopathy
TU	Trauma Unit
TXA	Tranexamic Acid
UK	United Kingdom
USA	Untied States of America
VHA	Viscoelasatic Haemostatic Assay
VTE	Venous Thromboembolism
WHO	World Health Organisation

CHAPTER 1

INTRODUCTION

1.1 Scope of the chapter

In this introductory chapter, the context in which pre-hospital medical providers make decisions about injured patients is described. The chapter starts with a description of the societal impact of trauma before progressing to the pathophysiology and contemporary management of haemorrhage, shock and Trauma induced coagulopathy (TIC). Next, theories of human decision making and how decisions can be improved with the use of appropriate support are discussed. A critical appraisal of the existing prediction models for major bleeding and trauma induced coagulopathy describes why so few models are in regular clinical use. The potential of machine learning in medicine is discussed, along with the strengths of a type of model called Bayesian Networks (BN). Finally, a BN developed to predict the risk of trauma induced coagulopathy in the Emergency Department is described.

1.2 Aim

The aim of his chapter is to establish the context in which clinical decision support models may augment decision making and improve outcomes after injury.

1.3 Trauma

1.3.1 The epidemiology of trauma

"Every day the lives of more than 14 000 people are cut short as a result of an injury"[1]

The World Health Organisation estimates that 5 million people die from physical injury each year and this number is predicted to increase.[2, 3] In 2016 injury was the 9th leading cause of death, by 2030 road traffic injuries alone will be the 7th highest cause of death. Injury encompasses acts of violence, road traffic collisions (RTCs), burns, drowning, falls, and poisonings.[4] Amongst these suicides, homicides and RTCs are the leading causes of injury and account for 50% of all injuries globally. The physical, mental, and financial costs of these injuries not only affect the injured individual but also have a detrimental, and sometimes devastating, impact on the families and communities affected. In part, this is because the majority of patients survive their injuries, but many survivors of trauma are left permanently disabled. As trauma disproportionally affects the young, injury causes many years of living with a disability.[5] This demand on the health care system is represented in **Figure 1**.

In high-income countries, trauma is often the leading cause of death in young people.[6] In England and Wales, trauma causes more deaths in the 15-44 year old age group than any other disease.[7] The Trauma and Audit Research Network (TARN) collects injury data from participating hospitals in the United Kingdom (UK). This data demonstrates that transfer of blunt force energy such as a road traffic collision causes the majority of injury leading to hospitalisation in the UK. Between 2008 and 2016, a blunt mechanism of injury was responsible for 96% of all TARN submissions. The remaining 4% are due to penetrating injury.[8]

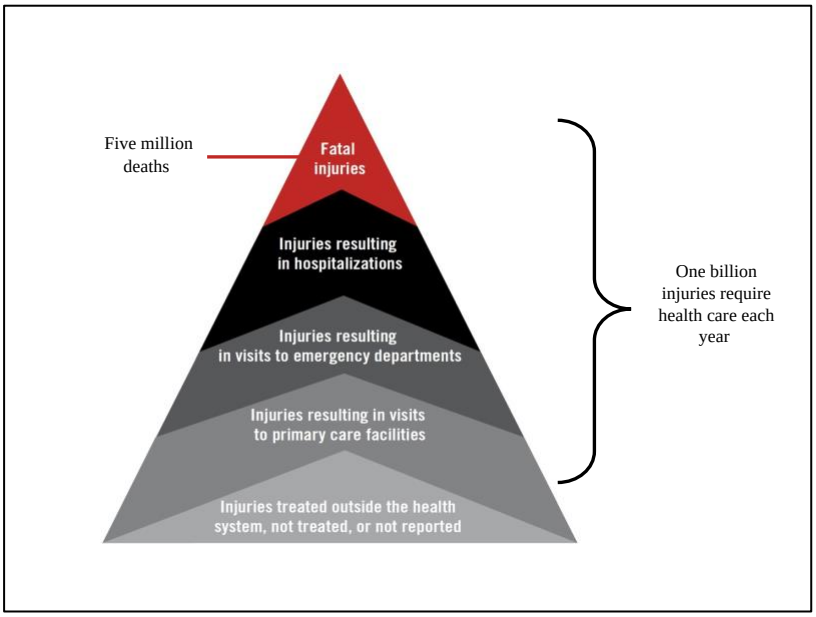

Figure 1: Graphical representation of the global demand on the health sector caused by injuries. Adapted from the World Health Organisation. Injuries and violence: the facts 2014.[1] Image reproduced under the cretive commons licence Licence: CC BY-NC-SA 3.0 IGO.

One notable trend in the demographics of trauma patients is an increasing number and proportion of elderly patients. In 2008, 22% of patients submitted to TARN were over 64 years old, but by 2016, this number had risen to 42% in the same hospitals. Over time TARN has engaged with a higher proportion of hospitals in the UK, such that the exact proportion of injured patients over 64 years old may even be closer to 50% in 2016.[8] This marked increase likely reflects both an ageing population and perhaps, more importantly, a heightened recognition of severe injury in the elderly. For example, the proportion of patients in the TARN database registered as injured by falling from less than two meters has grown at a similar rate.[9]

While a rise in elderly trauma represents a national trend, there are notable regional variations within the UK. In urban areas such as London, there is increasing penetrating injury due to interpersonal violence with knives.[10, 11] The Office for National Statistics recorded a 20% rise in offences involving the use of stabbing implements in the year ending in March 2017.[12] The peak incidence of knife injuries occurs in patients aged between 16-24 years old.[12]

It is vital to understand the epidemiology of trauma to provide appropriate systems of care to patients. For example, the mechanism of injury is related to the proportion of

patients requiring an operation. In a series of 160,000 patients submitted to TARN between 2014 and 2016, 43% of penetrating and 2% of blunt injured patients required urgent surgery.[13] Operations in this context are time-critical and delay to surgery increases the risk of death from haemorrhage. Therefore, regions with high proportions of penetrating injury need to ensure access to early surgery.

1.3.2 **Systems of trauma care**

The aim of a trauma system is to save lives and to return injured patients to a quality of life as close as possible to their pre-injury status. To achieve these aims requires intricate co-ordination of multiple resources. A trauma system is thus an integrated, and systematic structure which provides a multidisciplinary response to injured patients from the point of injury to rehabilitation.[14] Different configurations of trauma systems have developed across the world, varying by geography, technology, resources and politics.[15]

Regardless of the specific trauma system, inclusive systems provide the same four key phases: pre-hospital care, reception and resuscitation, definitive care and rehabilitation. For severely injured patients, care may start at the scene of injury or a patient may go directly to hospital without treatment. In-hospital care includes the emergency department, radiology, operating theatres, intensive care and the wards. Following discharge, patients may continue to access care from therapists, their primary care doctor and attend outpatient appointments at the trauma centre.

Established inclusive trauma systems concentrate patients and resources at a few high-volume trauma centres. The resultant high volume is associated with improvements in the delivery of care and outcomes for severely injured patients, including a reduction in mortality.[8, 16-18]

In this thesis, the research concerns two specific trauma systems: the civilian Trauma System in England and the deployable British Military Trauma System.

1.3.3 **The trauma systems in England**

The provision of trauma care changed in England in 2012. Following multiple reports of regional variability and deficiencies in the standards of trauma care,[19-21] the country is now organised into 17 Trauma Networks, containing 27 Major Trauma Centres. Each Trauma Network is a collaboration of services providing trauma care in a geographical area. These providers include pre-hospital services, the Major Trauma Centres (MTC), other hospitals receiving trauma admissions (known as Trauma Units), and rehabilitation services. MTCs provide specialised care for patients with major trauma. A patient is said to sustain major trauma when their injury is life-threatening or life-changing due to long-term disability.

The objective of a trauma system is to get the right patient, at the right hospital, at the right time, receiving the right care.[22] Trauma systems have developed "major trauma triage tools" to help differentiate patients at high risk of severe injuries. These tools are used by pre-hospital crews to help decide on the most appropriate hospital facility for each patient. In general, patients identified as high risk are taken directly to an MTC. Exceptions to this rule are if the transfer time exceeds 45 minutes or if the patient has an immediately life-threatening condition. In these instances, ambulances are directed to take patients to the nearest Trauma Unit (TU) for stabilisation. In circumstances where the facilities of the TU are insufficient to manage the patient's injuries, patients will be transferred to the MTC within their Trauma Network after initial assessment and resuscitation.[23] This is called a secondary transfer, which may also occur if an injured patient presents directly to a TU.

1.3.4 **The British Military Trauma System**

The Operational Care Pathway is military doctrine, prescribing the structure and delivery of the military trauma system.[24, 25] The UK Defence Medical Services (DMS) is the organisation coordinating and delivering care during training and operations. There are many similarities between the civilian and military trauma systems. The ultimate goal of both systems is to save lives and rehabilitate the injured. The four critical phases of care are the same in both systems: pre-hospital care, reception and resuscitation, definitive care and rehabilitation. The military trauma system is one

complete system, providing care from the point of wounding (Role 1) to rehabilitation (Role 4) often across several countries.

The military trauma system is organised to meet a series of key objectives. The 10/1/2/2 rule stipulates that within 10 minutes of injury, the casualty should receive enhanced first aid. This immediate care is initially delivered by military first-aiders and then by trained military medical personnel. Within 1 hour, the casualty should receive Enhanced Field Care often provided by a doctor or senior nurse; this is Role 1 care. Forward medical evacuation takes casualties from close to the point of wounding to forward surgical facilities known as Role 2. The forward surgical facilities provide a damage control surgery capability, ideally within one hour but within 2 hours at the longest. A patient may pass further along the in-country trauma system to Role 3, which is a more substantial medical treatment facility with higher capacity and capability. This final part of the 10/1/2/2 rule aims to deliver in-theatre surgery and diagnostics within 2 hours. Strategic casualty evacuation takes casualties from the Role 3 hospital and delivers them to the UK at the Royal Centre for Defence Medicine in Birmingham. Following hospital discharge onward rehabilitation is provided by the Defence Medical Rehabilitation Centre.

The last large-scale deployment of the British Military Trauma System ended in 2014 with the end of combat operations in Afghanistan.[26] As with results from civilian trauma systems[8, 17, 18] analysis of the probability of survival within the military trauma system saw improvements year-on-year. Throughout the conflict, the injury severity score, measured by the New Injury Severity Score (NISS), associated with a 50% chance of survival rose from a score of 33 to 60 (**Figure 2**).[27]

1.3.5 **The pre-hospital phase of care**

1.3.5.1 Delivery and organisation of pre-hospital care

Pre-hospital emergency care (PHEC) denotes care before hospital arrival. There are many similarities between Civilian and Military PHEC. One important common factor is the numerous types of care provider. In the civilian major trauma system, a patient may initially receive care from by-standers acting as good Samaritans. The ambulance

service is usually summoned via an emergency call to 999, and statutory pre-hospital care is delivered by ambulance trusts staffed by emergency technicians and paramedics. Supporting the ambulance response, Air Ambulance Services transport specially trained pre-hospital clinicians to the scene of the injury. These Helicopter Emergency Medical Services (HEMS) typically attend the most severely injured and time-critical cases. Their dual-role is to expedite delivery of experienced medical personnel to the scene and to convey patients to hospital as quickly as possible.

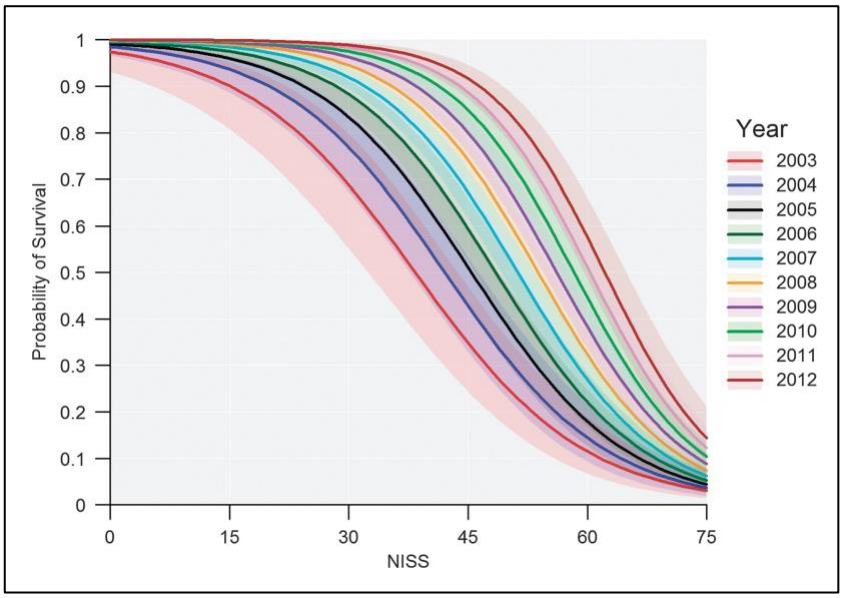

Figure 2: The predicted probability of survival by New Injury Severity Score for each year. Shaded regions indicate the 95% CIs for the predicted values obtained from a logistic regression model.[27]

In the military pre-hospital trauma system, immediate care is dependent on operational and tactical constraints. The level of care provided to a casualty is dependent on the threat posed by the tactical situation and the risk to the care providers' safety. Within the DMS, medical pre-hospital emergency care providers are distinguished by levels according to the NHS Skills for Health structure (**Table 1**).[28]

Table 1: Defence Medical Services levels of Pre-hospital Emergency Care

Level 7-8	Consultants and registrars with specialist pre-hospital training
Level 5-6	Any doctor, senior nurses and paramedics
Level 4	Combat medical technicians
Level 1-3	First aiders

Tasking of medical assets is usually reported by radio to a Patient Evacuation and Co-ordination Centre (PECC). Immediate response teams or Medical Emergency Response Teams (MERT) provide forward medical evacuation. The medical staffing and evacuation platform varies by operation. During the conflicts in Iraq and Afghanistan, a consultant in Anaesthesia or Emergency Medicine, a flight nurse and two paramedics staffed the MERT. These conflicts provided MERT in CH47 helicopters. By the end of the conflict in Afghanistan, the MERT was capable of delivering pre-hospital anaesthesia, blood transfusions and a limited repertoire of life-saving surgical interventions; all while simultaneously transporting the patient to hospital.[29] Future conflicts are predicted to be less amenable to this level of medical sophistication.[30]

Figure 3: Comparison of the space to provide medical care in two medically fitted helicopters. Left: Military; inside the Chinook (CH47), photograph from Greaves[26], Right: Civilian; inside the AW169 helicopter used by Air Ambulance Kent, Surrey and Sussex photograph authors' own

1.3.5.1 Opportunities and challenges of pre-hospital trauma care

PHEC provides a unique set of opportunities and challenges. Pre-hospital practitioners are the first step in the chain of survival for an injured patient. For some patients with life-threatening injuries, early intervention is the key to saving lives. In the military context application of tourniquets to bleeding extremities as part of PHEC leads to increased survival after blast injury.[31] The care delivered soon after injury sets the patient along a trajectory that can either mitigate or exacerbate the deleterious effects of trauma. Effective pre-hospital care may not only save lives for time-critical patients but also reduce the financial and biological cost of injury to the patient and the entire trauma system.

Pre-hospital medicine is often characterised by difficult decision-making; whether to stay on scene to perform medical interventions and thereby delay transport to definitive care, or to minimise interventions and move patients to definitive care as quickly as possible. Colloquially, this is known as the "stay and play" vs "scoop and run" predicament. Amongst well-developed healthcare settings, there are notable differences in the way pre-hospital care is organised. In part, these differences reflect differing attitudes to the effectiveness of pre-hospital interventions on patient outcome.

In the United States (US), very few pre-hospital systems incorporate physicians. Broadly, the US model prioritises rapid patient transport to hospital in both civilian and military systems of care. This approach is supported by a body of research, including articles demonstrating increased survival in patients moved rapidly to hospital by methods of transport without medical capabilities.[32, 33] In the UK, by contrast, Air Ambulance services provide medical care with physicians often present as part of the service. The senior decision-making provided by experienced pre-hospital doctors may help to improve survival by identifying and treating time-critical injuries. Determining which paradigm of pre-hospital medical care affords the best chance of survival is an area of ongoing research.[34, 35]

Dr John Holcomb, a US trauma surgeon with extensive military experience, recently described the need to move critical interventions such as non-compressible haemorrhage control forward of the hospital. Holcomb established that the median time to a haemorrhagic death in Houston, Texas, was 1.65 hours. Meanwhile, the median time to definitive surgical control of bleeding was 2.1 hours.[36] He argued that by placing specific haemorrhage control interventions forward of the hospital, a higher proportion of injured patients could be treated and may survive (**Figure 4**).

It is not possible to give a definitive verdict on whether "stay and play" or "scoop and run" is better. Most likely, the optimum strategy depends on the patient, their injuries and involves a combination of providing urgent care while transporting the patient expeditiously to a hospital. Deciding how much medical care should be provided pre-hospital requires a prediction by the treating clinician of the patient's specific risks and available therapies.

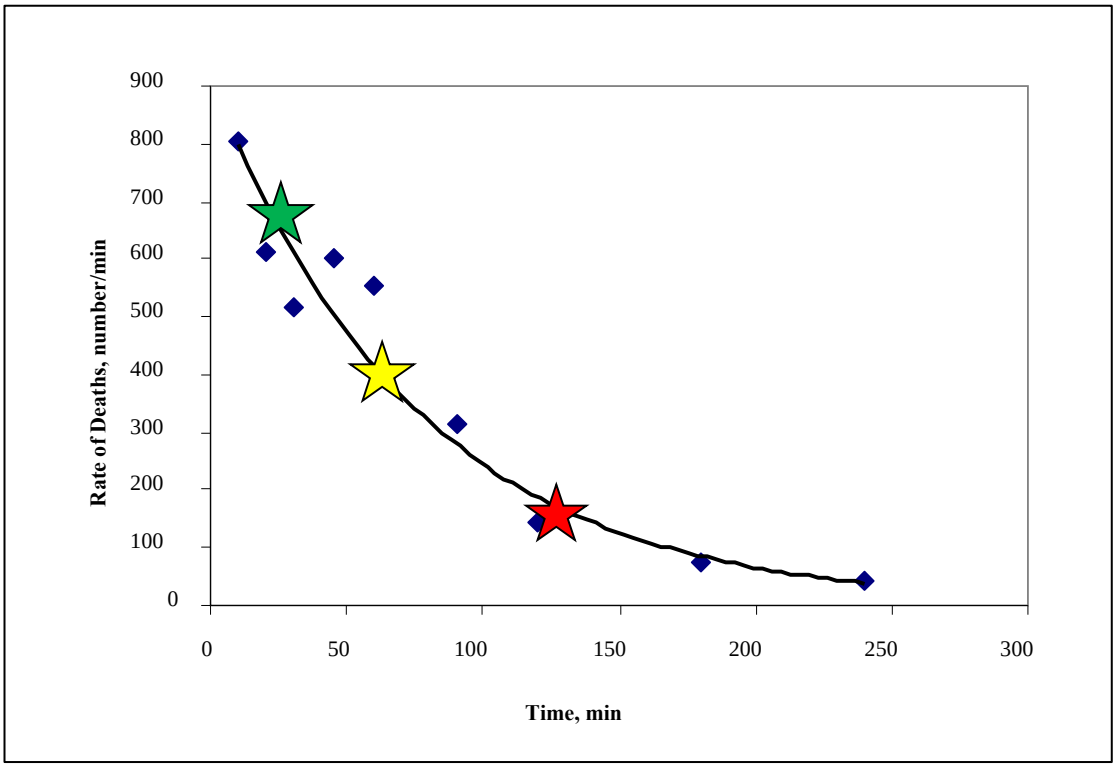

Figure 4: Relationship between the rate of death and time in 55,537 fatalities of motor vehicle collisions in the US. Modified from Holcomb[36]. Green star; time to deliver pre-hospital intervention, Yellow star; time to emergency department intervention, Red star; time to operating room definitive haemostasis.

1.4 Haemorrhage, shock and coagulopathy

1.4.1 Clinical significance

Each year 1.5 million people die from haemorrhage after traumatic injury.[37] Despite a concerted effort, bleeding remains the principal cause of preventable death in both civilian and military environments.[38, 39] Bleeding is responsible for half of the all of the *early* deaths from trauma, i.e. within the first 24 hours of injury. Globally, haemorrhage accounts for 75 million years of life lost per year, and those who survive the initial haemorrhagic insult have poor functional outcomes.[40, 41]

Severe blood loss causes haemorrhagic shock. This shock causes inadequate oxygen delivery to organs, cells and ultimately mitochondria.[42] If bleeding continues unchecked, the result is death. The median time to death from haemorrhage is

approximately two hours.[43, 44] The magnitude and duration of inadequate cellular perfusion may be related to the subsequent pathological effects.[45] In particular, shock is associated with multiple organ dysfunction syndrome (MODS)[46] and trauma induced coagulopathy (TIC).[47]

TIC is a multidimensional failure of the coagulation system resulting from a combination of tissue injury and shock.[48-50] Roughly a quarter of severely injured trauma patients have laboratory-based evidence of TIC and [51] death occurs in 50% of those who develop TIC.[52] A high burden of morbidity is also associated with TIC such as exacerbation of bleeding, increased multiple organ dysfunction, prolonged intensive care stays and increased thromboembolic events.[53-56 57]

1.4.2 **Pathophysiology**

Shock is a state in which the oxygen demands of respiring tissues are greater than oxygen supply.[37] At the cellular level, inadequate oxygen delivery leads to anaerobic metabolism, accumulation of cellular by-products, the release of damage associated molecular patterns (DAMPs) and activation of inflammatory pathways. Once cells have consumed their stored energy (adenosine triphosphate), homeostasis fails, and cell death occurs. At a tissue level, critical hypoperfusion of organs results in organ dysfunction. For the brain and myocardium, dysfunction results in cerebral anoxia, arrhythmias, and death within a few minutes.[37]

The role of the vascular endothelium is key in linking hypoperfusion with coagulopathy. In shock states, there is endothelial cell activation and shedding of the endothelial glycocalyx.[58, 59] This *endotheliopathy* is associated with coagulopathy, MODS and mortality.[60, 61] Dissolution of the protective glycocalyx layer liberates heparin and exposes blood constituents to the endothelial cells and surface linked proteins such as thrombomodulin. A result of endotheliopathy is altered coagulation haemostasis.[61]

TIC is the clinical manifestation of an endogenous multi-factorial and multi-mechanistic failure of the coagulation system, exacerbated by improper resuscitation techniques.[62] Brohi describes Acute Traumatic Coagulopathy (ATC) as an early endogenous process mediated by the protein C pathway in response to tissue injury and hypoperfusion.[63, 64]

Dilution, hypothermia and acidosis all cause an iatrogenic coagulopathy.[65] However, there are a wide range of alternative hypothesized mechanisms for TIC. The range of mechanisms is likely due to global research efforts to date selecting a multitude of coagulopathy phenotypes and studying patients which vary in their presentation.[62]

The evidence supports the need for both tissue injury and tissue hypoperfusion to cause TIC. There is no widespread agreement on the precise mechanisms, and the following hypotheses have all been implicated: increased activated protein C[52, 54, 63, 66, 67], clotting factor depletion[68, 69], abnormalities in fibrin generation[70], DAMPS and cross-talk with inflammation[71, 72], increased fibrinolysis and fibrinogenolysis[73, 74], a controversial fibrinolysis shut down theory[75, 76], platelet dysfunction with loss of platelet aggregation[77-80], endotheliopathy[81, 82] and disseminated intravascular coagulopathy with a fibrinolytic phenotype.[83-88]

Trauma induces multiple biochemical and physiologic changes, and despite numerous studies reporting differences in coagulation parameters between trauma patients and uninjured controls, it is unclear whether some of these differences may be "normal" after trauma. Fundamentally, it is not clear whether ATC is an aberrant pathological response or an evolutionary selected adaptive response. The intended benefit of this adaptive response may be to maintain tissue perfusion during haemorrhagic shock and mitigate the resultant multiple organ failure.[62]

Whether this process is adaptive or maladaptive remains an ongoing debate. Some believe there is evolutionary benefit from processes which may restore perfusion in shocked states and avoid organ failure. Whereas, others believe it is a purely pathological phenomenon which does not have an evolutionary selection benefit.[62] Regardless, patients that develop TIC are at a high risk of death and need to be identified early and managed appropriately.

1.4.3 Why is it potentially useful to predict patients at risk of haemorrhage, shock and coagulopathy pre-hospital?

Good judgement and sound decision-making are essential to safe and effective clinical practice. Early recognition of patients with substantial haemorrhage, shock or

coagulopathy after injury is thus critical to improving patient outcome. There are multiple decisions which are informed by an awareness of the patient's physiological response to injury . Once a pre-hospital clinician identifies a patient as high risk, they can choose to deliver a range of interventions aimed at repaying the oxygen debt and preserving homeostasis.

The concept of minimising the time to delivery of a critical intervention is well recognised throughout medicine. Studies in stroke[89], acute myocardial infarction and sepsis[90] demonstrate improvement in outcome with shorter intervals to therapy. In trauma, treatment delays are associated with increased mortality. This relationship is observed in civilian trauma patients undergoing laparotomy[91], combat casualties arriving at a medical facility[92, 93], and time to receiving blood products.[94] In bleeding patients every minute delay in receiving blood products is associated with a 5% increase in the odds of death.[94]

The prompt availability of blood products is an essential component of 'Damage Control' interventions. However, there are many logistical steps in preparing blood components for transfusion. Hospitals have Major Haemorrhage protocols designed to minimise these delays in emergencies. Nevertheless, effective activation of these protocols depends on early identification of bleeding patients, which is difficult.

1.4.3.1 Why specifically predict trauma induced coagulopathy pre-hospital

Patients that develop TIC have poorer outcomes than those that maintain normal coagulation (see **1.4.1**). Injured patients at high-risk of poor outcome should be managed according to the principles of damage control resuscitation.[95] This includes the prioritisation of control of bleeding, delivery of oxygen to the cells, and maintaining haemostatic resuscitation.[96] Effective DCR results in earlier correction of TIC, reduction in the total volume of blood products used and decreased exposure to potential harm from the blood products.[97, 98]

By providing clinician users with accurate risk estimates (probabilities) at the point of care, a TIC prediction model aims to improve their situational awareness and thereby support earlier, better informed and personalized treatment decisions. Augmenting the

early recognition of patients with TIC by AI-systems therefore allows clinicians to co-ordinate the appropriate trauma system response. For high-risk patients pre-hospital, the right response includes appropriate triage of the patient to a specialised major trauma centre and mobilisation of resources within the receiving hospital. These steps reduce the delays to treatment and may improve patient and system outcomes.

1.4.4 **Making a diagnosis**

1.4.4.1 Contemporary diagnosis of haemorrhage and hypovolaemic shock

The diagnosis of haemorrhage is fundamental to the management of an injured patient. Quantifying the degree of blood loss is a clinical consideration made for every trauma patient (CHAPTER 2). Yet making an accurate diagnosis when bleeding is occult, remains challenging. Diagnosis of bleeding is especially challenging in environments with minimal diagnostic aids such as pre-hospital care. Clinicians are liable to both over-triage and under-triage. In a multicentre study in the US, trauma surgeons in the Emergency Department only identified one-third of patients requiring the highest level of blood transfusion.[99]

Traditional courses like the American College of Surgeons' Advanced Trauma Life Support (ATLS) Course, teach doctors to recognise the degree of blood loss using vital signs (**Table 2**).[100] [101] The classical theory suggests that steady derangement in vital signs correlates with the degree of blood loss. In reality, the relationship between the degree of blood loss and vital sign derangement is far more variable and less predictable. Thus, the diagnosis of hypovolaemic shock using ATLS's vital sign thresholds is inaccurate. Analysis of large databases of trauma patients demonstrates the inaccuracy.[102-106] In one prospective study, systolic blood pressure had a sensitivity to detect haemorrhage of only 26%.[107]

Part of the fallibility of "simple" vital sign thresholds is the individual variation in human physiology. Multiple homeostatic responses are trigged during shock and include an initial phase of tachycardia and vasoconstriction followed by a second phase of reflex bradycardia and vasodilation.[108] Humans have evolved multiple mechanisms to compensate for blood loss and maintain homeostasis for as long as possible. When these

mechanisms are overwhelmed, decompensation occurs, and the vital signs become precipitously deranged as blood flow plummets.[109] The individual variation in the ability to compensate for blood loss means that flow cannot be reliably inferred from traditional vital sign measurements.

Table 2: ATLS Classes of haemorrhagic shock

	I	II	III	IV
Blood Loss* (mL)	Up to 750	750-1500	1500-2000	>2000
Blood Loss (%)	Up to 15	15-30	30-40	>40
Pulse Rate (per minute)	<100	100-120	120-140	>140
Blood pressure	Normal	Normal	Decreased	Decreased
Pulse pressure	Normal	Decreased	Decreased	Decreased
Respiratory Rate (per minute)	14-20	20-30	30-40	>35
Urine output (mL/hour)	>30	20-30	5-15	Negligible
Mental status	Slightly anxious	Mildly Anxious	Anxious, confused	Confused, lethargic

*Blood loss volume for a male patient weighing 70kg

Scrutinising the changes over time of a patient's heart rate does not appear to improve diagnostic accuracy either. A study of three pooled datasets reported that 41% of trauma patients with major haemorrhage did not respond with a tachycardia. In this study, almost all the bleeding patients became hypotensive within 30 minutes of monitoring. Contrary to classical teaching, on the onset of hypotension had no clear correlation with a trend of increasing tachycardia.[106]

To determine the degree of blood loss, England's National Institute for Health and Clinical Excellence (NICE) guidelines recommend assessing a patient's response to a fluid bolus.[110] Following a fluid bolus, a patient's haemodynamic status can be said to either respond, not-respond, or transiently respond. The inference is drawn that non-responders and transient responders are actively bleeding.[111]

Expert approaches to make a diagnosis of haemorrhage combine a range of clinical and point of care diagnostic tests.[112] In Chapter two, these strategies are thoroughly discussed. In brief, clinicians combine information from the mechanism of injury to

search for the site of potential bleeding, assess subtle patient parameters such as the patient's higher cognitive function, and qualitative signs such as pallor, clamminess and venous guttering. Pre-hospital these may be the only methods used to diagnose blood loss. Some pre-hospital services use point of care diagnostics like blood lactate.[113, 114] Once the patient arrives in hospital, a range of additional tests including imaging technique such as X-ray, Ultrasound and Computed Tomography, are available to support decision making.

Two recent technological advances show promise in improving the accuracy of the classical vital signs to haemorrhagic shock. Non-invasive microcirculatory imaging quantifies blood flow within the microcirculation. This technique has demonstrated that two patients with the same systemic systolic blood pressure may have significantly different perfusion of their microvasculature.[115, 116] Another recent technology-enabled device is the Compensatory Reserve Index (CRI). The CRI analyses plethysmography waveforms using a machine learnt algorithm trained on healthy volunteers with simulated hypovolaemia. CRI is designed to quantify the degree of cardiovascular compensation at a given time by measuring changes in the pattern of blood flow waveforms. As blood loss progresses vasoconstriction increases and this can be observed with the CRI device. Still, in a development phase, early research suggests the CRI measurements are more sensitive to changes in circulating volume than the classical vital signs.[117, 118] Both microcirculatory imaging and the CRI require further investigation to establish their clinical utility.

1.4.4.2 Contemporary diagnosis of trauma induced coagulopathy

Traditionally TIC was diagnosed if conventional laboratory tests exceed given thresholds. The literature contains multiple thresholds to classify TIC (see **7.5.4**). Some of the most frequently used cut-offs included 50% prolongation of Prothrombin Time (PT) or partial thromboplastin time (PTT).[55] This definition of coagulopathy came from the guidance on coagulopathy of any cause. There are numerous problems with this approach. First, PT and aPPT are performed on platelet-poor plasma and require 30-60 minutes to process. The cell-based model of haemostasis emphasises the importance of the platelet.[119] For rapidly bleeding patients results do not reflect the contemporary coagulation function[120], and the endpoint of the test is the start of clot formation.

In an important study in 2010, Frith *et al.* noted that the threshold values used to classify patients with TIC were arbitrary, and the relationship between cut-off value and clinical outcome was unknown.[121] In a multicentre study of over 3500 patients, the authors noted that patients arriving in hospital with a prothrombin time ratio (PTr) > 1.2 had significantly higher mortality and transfusion requirements.[121] In doing so, the authors defined a clinically meaningful threshold value.

In a study of a London MTC, the median time for PT results to be available to clinicians was 78 minutes (IQR 62-103).[122] Point of care coagulation assays that measure INR are often used in the community to guide anticoagulation dosing. These devices offered an attractive method to reduce the time to clotting results. Unfortunately, the accuracy of these devices to detect TIC has been deemed insufficient.[122] With studies suggesting a lack of correlation with laboratory results in the presence of haemorrhage and anaemia.[122, 123]

Viscoelastic haemostatic assays (VHA) also known as functional coagulation analysers such as rotational thromboelastometry (ROTEM®, Tern International GmbH, Munich, Germany) and thromboelastography (TEG® Hemoscope, Niles, IL), have several advantages over conventional coagulation tests (CCT).[124] These tests can be used to provide relatively fast point of care coagulation testing. Additionally, VHA are whole blood assays and with added activators and inhibitors can assess different stages of clot homeostasis. Using ROTEM, Davenport *et al.*, demonstrated that the primary functional disorder of TIC was a loss of clot strength.[122] In a prospective study of 300 trauma patients, an ExTEM clot amplitude at 5 minutes less than 36mm was more sensitive than PTr >1.2 for identifying patients with coagulopathy and receiving a massive blood transfusion.[122, 125] Other studies have used VHA to establish diagnostic thresholds for the hyperfibrinolysis of TIC.[126] Although the sensitivity of VHA assays to detect clinically important fibrinolysis is low.

VHA diagnostics are improving in usability and portability.[127] However, there are still no VHA devices in use for the pre-hospital diagnosis of TIC. Several novel devices are in development that aim to provide accurate and portable pre-hospital testing of coagulation.[128-130] There are also ongoing research efforts to establish the reliability of

current devices amongst the additional environmental challenges encountered pre-hospital; such as engine vibration on helicopters.[127 131, 132]

In summary, making a diagnosis of TIC at the time treatment decisions have to be made is difficult. Laboratory results from conventional coagulation tests describe only some of the clotting abnormalities and are not available early enough to guide therapy.[122, 133] Thromboelastometry is more useful but requires sensitive hardware that makes it unsuited to use in austere or mobile care settings. Current practice, therefore, relies on blind, empirical protocols or expert clinical opinion, and the ability to accurately and objectively tailor treatment decisions to individual risks is awaited.

1.4.5 **Pre-hospital management of haemorrhage, shock and coagulopathy**

Damage control resuscitation (DCR) is the overarching treatment strategy designed to limit the pathological effects of shock and coagulopathy.[134-136] The central tenets of modern DCR are prioritisation of early haemorrhage control while maintaining haemostatic resuscitation.[111] Guidelines from the Eastern Association for the Surgery of Trauma and NICE expands these principles into further objectives: avoid hypothermia, minimise blood loss with early haemorrhage control during transport and initial evaluation, target low-normal blood pressure before definitive haemostasis, minimise crystalloid use, use major haemorrhage protocols, avoid delays to surgical or angiographic haemostatic, transfuse blood products that optimise haemostasis, obtain functional laboratory measures of coagulation and give pharmacologic adjuncts to promote haemostasis.[110, 137] The use of DCR is associated with improved outcomes in appropriately selected patients.[137]

DCR starts pre-hospital and often continues several days following the injury. DCR can be considered in four phases: pre-surgery, initial surgery, restoration of physiology and definitive surgery. In the physiologically deranged patient surgery is limited to haemorrhage control and contamination containment with abbreviated surgical techniques.[134, 138-140] After initial damage control surgery (DCS) the aim is to repay the oxygen debt and return physiological haemostasis before definitive surgery. In the pre-surgery phase improved survival in high risk patients is associated with control of

haemorrhage, early haemostatic blood transfusion and rapid transport to an appropriate facility. [31, 36, 141-143]

This thesis is about the decisions made early after injury within the pre-surgery phase of DCR. The following chapters explore how clinicians recognise patients that require DCR and whether these decisions can be supported with decision support aids. The recognition of appropriate patients is the first step in instigating DCR.

1.4.5.1 Pre-hospital blood transfusion

Pre-hospital transfusion has its origins in World War 1.[144, 145] Pre-hospital transfusions continued to be used by military doctors in conflicts such as Vietnam.[146] However, there was no widespread use of pre-hospital blood transfusion in civilian practice until renewed recognition that earlier transfusion of bleeding patients may be related to improved patient outcomes. The British Military was amongst the first to re-introduce pre-hospital blood transfusion.[147] The British Defence Medical Services added a blood transfusion capability to the Medical Emergency Response Team (MERT) during the conflict in Afghanistan in July 2008. Civilian practice followed and in March 2012, London's Air Ambulance started to carry blood pre-hospital. Each pre-hospital service developed its own protocolised transfusion triggers.[147, 148]

The hypothesis that *pre-hospital* blood or blood products leads to improved patient survival is an area of active research.[149] Much of the current evidence, from both military and civilian settings, is retrospective and while largely supportive of the hypothesis, has unavoidable methodological risk of bias.[93, 150-157] Recently two randomised controlled trials sponsored by the Department of Defense reported their findings with conflicting results. PAMPer[158] and COMBAT[159] both ostensibly asked the same question: Did two units of plasma given pre-hospital reduce mortality? PAMPer demonstrated lower mortality in the group receiving plasma vs standard care in 501 trauma patients transported by air (23% vs 33%; p=0.03).[142] Whereas, COMBAT saw no difference in mortality between those treated with plasma and those randomised to normal saline and transported by ground ambulance (15% vs 10%; p=0.37).[160] Several causes for the observed difference have recently been proposed and include differences in patient selection, mechanism of injury and time to definitive treatment.[161-164] In England, the

ongoing Re-PHILL trial also aims to establish the relationship between pre-hospital blood and patient outcomes.[165]

The mixed results of well-conducted clinical trials have not changed the practice of many pre-hospital services who continue to transfuse blood and blood products in critically shocked patients. Blood transfusion is not risk-free. Risks include infections transmitted by the transfusion, immune-mediated complications and transfusion-related cardiac overload. As a result of these risks, logistic and financial costs, transfusion is restricted to patients most likely to benefit from the treatment.

1.4.5.2 Major Haemorrhage Protocols

A Major Haemorrhage Protocol (MHP) is a key facet of the haemostatic principle of DCR, and its use is associated with improved survival.[96, 98, 166-173] An MHP is a multi-modal package of care that provides standardised elements of medical treatment to patients with life-threatening haemorrhage after injury. MHPs are designed to streamline balanced blood product delivery to bleeding patients and allow DCR to begin before the availability of laboratory results.

An MHP includes the activation criteria to trigger the protocol, the major transfusion protocol (which is a predefined ratio of blood components for transfusion), haemostatic adjuncts, instructions for the reversal of anticoagulants and prevention of hypothermia, and team communication strategies. The purpose of combining and protocolising these elements of care is to ensure critical components are brought together rapidly and reproducibly. For example, it takes 30 minutes to thaw Fresh Frozen Plasma (FFP); a vital component of a blood transfusion in trauma. An MHP facilitates transfusion of FFP rapidly bypassing multiple routine steps. The rate-limiting steps for transfusion of FFP include taking the blood sample, sending it to the laboratory for coagulation testing, interpretation of the result by a clinician, requesting FFP, thawing FFP and finally delivery and administration to the patient. The delays associated with this process lead to clinical deterioration and could be fatal.

The components and the triggers to activate an MHP vary depending on the institution.[174] It is the triggers for MHP activation that are of most relevance to the

research within this thesis. Current triggers for activation usually consist of three main criteria: vital signs, laboratory data, and physician discretion.[175] NICE guidelines recommend the physician clinically assess the extent of traumatic haemorrhage using a combination of the mechanism of injury, patient physiology, anatomical injury pattern and the patient's response to initial resuscitation. These elements and how expert clinicians combine these factors are explored in chapter 2.

In inclusive trauma systems, pre-hospital clinicians can activate the receiving major trauma centre's major haemorrhage protocol before the patient arrives at the hospital. When this happens, blood is made ready for transfusion, senior decision-makers assemble to receive the patient and radiology, and operating theatres are put on standby. Inappropriate activation wastes finite resources, has a financial cost and potentially draws resources and personal away from other duties.

1.4.5.3 Tranexamic Acid

Tranexamic Acid (TXA) is a synthetic amino acid derivative that inhibits plasminogen activation by blocking the lysine-binding site and so inhibits fibrinolysis. The use of TXA in trauma has been demonstrated to reduce mortality when given to patient suspected of bleeding early after injury in both civilian and military populations.[176-180] In the large multi-centre CRASH-2 trial, all-cause mortality and death from haemorrhage were significantly reduced in patients that received TXA (Relative Risk 0.91 and 0.85 respectively). Some argue that the beneficial effects of TXA's inhibition of plasmin are additional to reduced fibrinolysis. Plasmin has pro-inflammatory activity. Benefit from TXA may be derived from its anti-inflammatory effects and subsequent reduction in multiple organ dysfunction.[181, 182]. Whatever the precise mechanism of benefit, there is a clear signal that giving the drug earlier increases its effectiveness.[179, 183] The key to early administration is giving the drug in the pre-hospital phase. [184, 185]

Theoretical concern about increased thromboembolic events with the use of TXA does not appear to be borne out in large clinical trials.[186] Despite this evidence, many clinicians remain concerned on a theoretical and mechanistic basis. These clinicians believe in patients who do not have evidence of hyperfibrinolysis, TXA must increase

the risk of thromboembolism.[187] Proponents of this argument advocate reserving TXA for only those patients with diagnostic evidence of hyperfibrinolysis.[188]

1.4.5.4 Summary: management of bleeding and coagulopathy in pre-hospital trauma care

Early identification of patients who require pre-hospital damage control resuscitation remains challenging. The pre-hospital environment provides few diagnostic adjuncts and short times from injury to assessment. Yet trauma systems that can provide early intervention stand poised to achieve some of the most significant gains for patients.

Analysis of current practice suggests that clinicians both over-triage and under-triage.[99] In doing so, clinicians expose some patients to unnecessary risks and deprive others that stand to benefit from treatment. Situations arise in which it is difficult for the pre-hospital doctor to assess a patient's intravascular volume status, their bleeding rate, the duration of their bleeding and hence the urgency of treatment.[101] Consequently, clinicians are forced to make high-stakes decisions under considerable time pressure with incomplete information. These are difficult decisions.

1.5 Decision making and decision support

1.5.1 Decision making

There are multiple models to explain how humans make decisions. Early theorists believed that decision-making required the generation of two or more alternative options. This was followed by rational comparison to select the option with the highest utility. Modern theorists reject these earlier beliefs. Modern theories such as Kahneman and Tversky's Heuristics and Biases (HB)[189] and Klein's Naturalistic Decision Making (NDM)[190] provide differing explanations about how we make decisions.

Kahneman and Tversky's Nobel prize-winning theory is rooted in behavioural economics and supported by laboratory experiments. The HB theory assists in the understanding of decision making by considering that two metaphorical 'modes' of

thinking occur in everyday decisions: System 1 and System 2. System 1 thinking is the intuitive, rapid decision making used for most of the decisions we make. It is 'fast and frugal,' uses heuristics and is prone to reproducible errors and biases. System 2 thinking describes the slower, logical and more demanding analytical processes we use when confronted with a difficult problem. System 2 thinking is slower, taxing and generally avoided if a System 1 response is sufficient. These modes do not operate in isolation and many decisions are made from a blended process. Nevertheless, the two system model is useful in understanding decision making and in generating hypotheses about predictable errors in human decision making.[191]

1.5.1.1 Naturalistic Decision Making

Klein's NDM theory originates from the observation of master chess players and the realization that that experts perform better in real world scenarios than might be predicted by the HB theory.[192, 193] Supporting data is collected from the real world observation of professionals in the military, emergency services and safety critical industries. Like pre-hospital clinicians these professionals, make high stakes decisions characterised by uncertainty, dynamism, competing goals and time constraints. The application of NDM theory has informed several industries including influencing military training doctrine. In part, NDM's application focuses on decision requirements and the development of technology to support decision making.[194]

1.5.1.2 Recognition-Primed Decision Model

Central to NDM is the concept of a Recognition-Primed Decision (RPD). RPD describes how an expert's experience is used to form patterns that influence decisions.[194] An expert will approach a familiar situation and recognise patterns. If the pattern is familiar the expert typically selects a successful action based on their prior experience. Recognition involves assessing relevant cues, expected outcomes, plausible goals, and selection of the typical action to achieve the desired goal. (**Figure 5**)

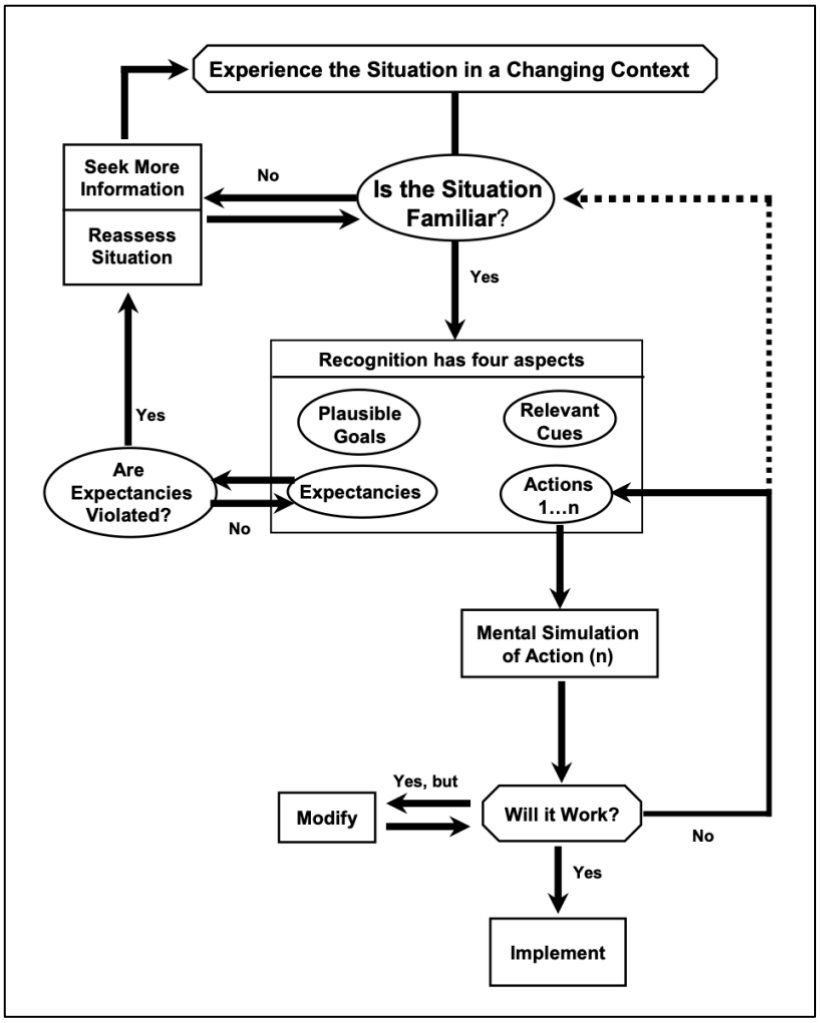

Figure 5: Schematic of recognition-primed decision making[195]

NDM theory includes the rapid mental simulation of a potential plan. This act of simulation provides the expert with an assessment of the likelihood of the desired goal being achieved with the selected action plan. If necessary, the plan can be modified, or another plan considered before it is implemented. Where no action plan satisfies the mental simulation step (dashed line **Figure 5**) the decision maker re-considers whether the situation is familiar. If it is not familiar, more information is sought, and the situation is re-assessed until a decision can be made. Klein describes the RPD model as a blend of intuition (HB's System 1) and analysis (System 2).[194]

Klein demonstrated that fire fighters faced with challenging scenarios used RPD strategies 80% to 90% of the time.[194] This observation has recently been replicated in healthcare. In a study of anaesthetists faced with difficult airway management situations,

91% of the decisions were recognition primed.[196] In this study of airway decisions anaesthetists characteristically made a direct link between familiar cues and action generation. The first action the anaesthetists considered was usually the action they implemented.[197]

To reinforce pattern recognition 'double loop learning' can be used. Double loop learning is a process where past assumptions are questioned by reflecting on a previous decision and then retesting that decision.[198] In healthcare, medical simulation training utilises the double loop learning concept. Trained personnel that have engaged in double loop learning are thought to be better *primed* to make the decision.

1.5.1.3 Example of Recognition-Primed Decision making in Trauma

The following example is used to illustrate the application of RPD to pre-hospital trauma decision making. The example describes the processes of cue perception, goal setting, expectations considered and action selection in the management of a patient with a tension pneumothorax after a fall (**Figure 6**). The example is simplified for clarity and appears sequential, but it is important to note RPD emphases the nearly simultaneous assessment of actions.

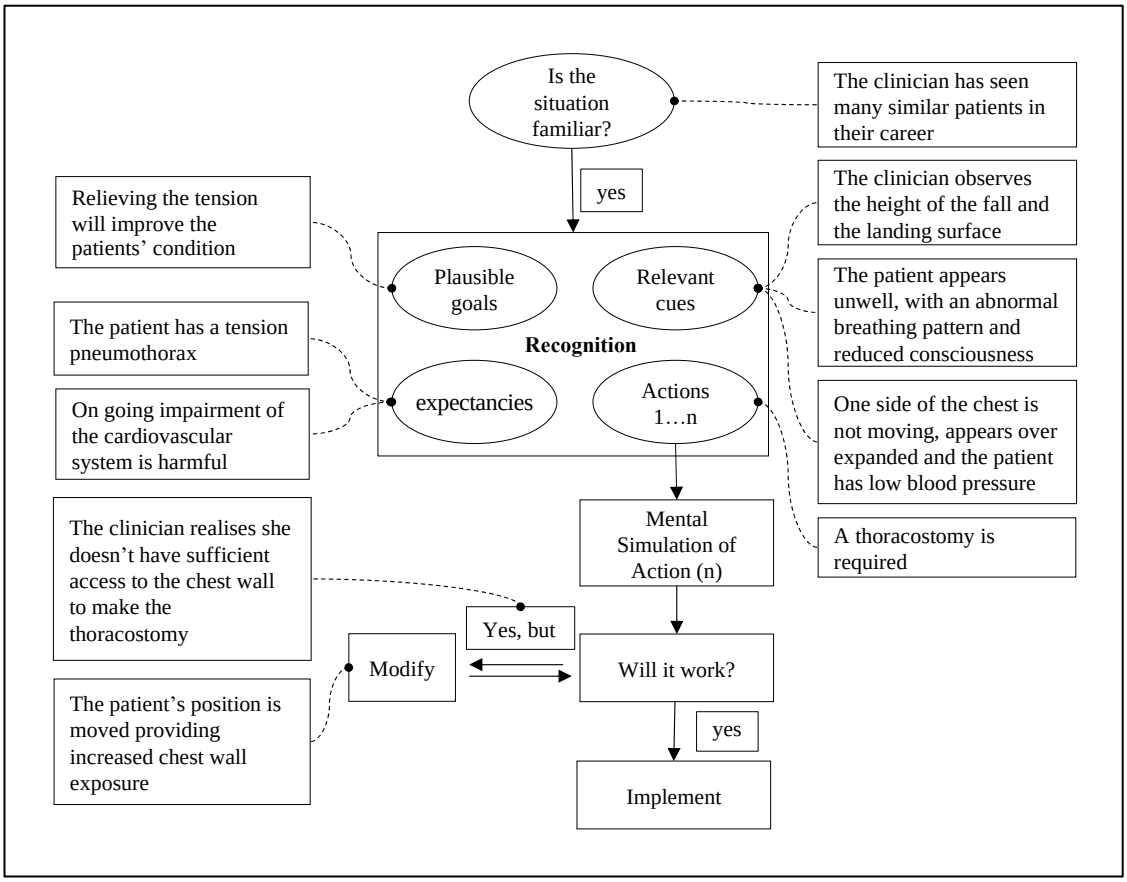

Figure 6: An example of Recognition-Primed Decision making in the management of a tension pneumothorax. Adapted from Klein.[195]

1.5.1.4 Heuristics & Biases and Naturalistic Decision Making: similar but different

HB and NDM share the belief that intuitive judgements are automatic and effortless. However, HB is sceptical of the value of intuitive judgment whereas NDM embraces expert intuition. The key difference is their stance on the origin of intuition. NDM attempts to explore intuition by eliciting the cues that are apparent to the experts (and elusive to novices). Whereas, HB research has focused on the use of simplifying heuristics as the basis of intuition and the tendency to systematic bias.

HB researchers often compare expert decision making to algorithms and advocate for the replacement of human judgement with formal models. The origins of this approach started with Paul Meehl in the 1950s.[199] NDM researchers are typically sceptical of attempts to formalise decision making with universal structures on complex

problems.[200] The application of NDM involves elicitation of key decision making cues and dissemination of those cues.

Both Kahneman and Klein agree that the quality of an intuitive decision is dependent on the ability of the decision maker to learn the regularity of their environment.[200] Where an environment is too uncertain to learn how cues relate to outcomes, decision making has a higher probability of being flawed.

Both HB and NDM theories provide insights into how decisions are made by pre-hospital clinicians. The theories also suggest where and how decision support may alter decision making. This is key when considering the application of decision support in medicine.

1.5.2 **Decision making under uncertainty in medicine**

Classical diagnostic decision making in medicine follows hypothetico-deductive reasoning. Clinicians start with a broad range of potential diagnoses. They update their hypothesis and narrow down on the probability of the patient having a given diagnosis by using information from the history, examination and investigations.[201] As discussed above most medical practice is not laboriously analytical (System 2) as classically described, but rather effortless (RPD / System 1). Yet, most medical error is believed to stem from errors of decision making.[202]

Uncertainty is a key feature of difficult decisions. Uncertainty increases the likelihood of a medical error. When clinicians make high-stakes decisions with uncertain information, under time or emotional pressure, the resultant errors can have devastating effects on patient outcomes.

Uncertainty describes a situation in which it is not possible to quantify the probabilities of alternative outcomes.[203] Quantification of the probabilities of alternative outcomes provides an assessment of risk. Moving from uncertainty to risk allows clinicians to make analytical decisions on prognosis and future management. For example, when a pre-hospital clinician is on their way to a patient, they have little useful information. Initially, there are too many unknowns to make a diagnosis; thus, a diagnosis of

pneumothorax for patient that has fallen is uncertain. When the clinician arrives at the patient's side, she rapidly gathers information and starts to recognise the patterns. With more information she is able to quantify the probability that the patient has a pneumothorax. At that point that risk can be used to estimate the benefit of a given intervention such as a thoracostomy.

Moving from uncertainty to risk is possible when informative evidence is gathered and synthesised. Sometimes the evidence may be conflicting, or the relative importance of a particular factor is difficult to put into context of the other information. In these scenarios, it is not clear how best to combine the evidence. The result is marked individual variation as personal biases influence assessment.[200] In these situations, an unbiased process which accurately combines the various inputs to produce an overall summary of risk may assist the decision-maker.

1.5.3 **Clinical decision support tools**

Clinical decision support tools (CDST) combine multiple inputs and produce an output summarising the influence of the inputs. The aim is to provide clinicians with information to make informed and improved decisions.[204] At the heart of a CDST is a mathematical function or model whose purpose is to combine the inputs in an appropriately weighted manner.[205] To define the weights of the inputs, algorithms are usually trained on prior data.

Simple clinical scoring systems, often derived from regression methods, are the most frequently used CDST in medicine. Examples of simple clinical scores include the Apgar score[206] and the Nottingham Prognostic Index[207]. However, conditions with multiple influential variables require complex models, which offer more precise quantification of an outcome than simple clinical scoring systems. Recent advances in computational methods provide the opportunity to develop precise models that allow risk, and subsequent treatment decisions, to be tailored to the individual patient.[208] A credible CDST must include an interface that is useable in the relevant clinical setting, and present results in a way that can be interpreted by the clinician. It should equip the user with the facility to perform evidenced-based analysis of individual patients,[209] improving decision accuracy to that exhibited by the most experienced practitioner. The

CDST should inform therapeutic decision-making at critical points in the treatment pathway and should also enhance overall clinical situational awareness. By supporting the correct interpretation of the clinical situation, decisions on the most appropriate action have a much higher chance of being correct too.[210]

1.5.3.1 Prognostic model research

Prognostic models combine multiple predictors to provide a likelihood of a patient having a particular outcome (endpoint) from a given state of health (start point) over a specific time interval.[205] Prognostic models ought to undergo a rigorous process of research before they can be safely implemented in healthcare.[211, 212] The purpose of the development process is to demonstrate a model's accuracy, generality and effectiveness (**Figure 7**).[213]

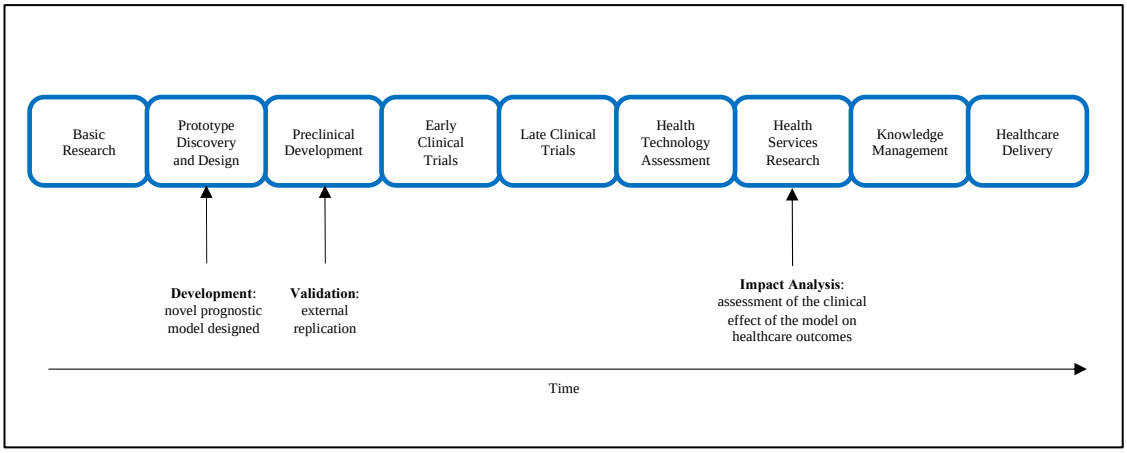

Figure 7: Position of key steps in model development along translational pathways. Adapted from Cooksey report and Steyerberg[205] (made available for use and re-use through the Open Government License doi:10.1371/journal.pmed.1001381.g003)

The first step is model development.[214] In development, an important and difficult clinical decision is identified along with the identification of the prognostic inputs. Next follows an initial training phase where the model learns the relationships between the variables using a dataset of training patients. Next, the performance of the model is quantified in the development set of patients; internal validation. Key performance metrics include discrimination and calibration.[215] A model that performs well in internal validation should undergo external validation. External validation tests whether the result of the model can be generalised. The patients in the external validation dataset

should ideally be fully independent in time and place to the development population.[216] Relatively few models in healthcare get this far in the development process.[217]

Only when changes in clinical management are made based on a model's prediction can the model be said to have influenced patient outcomes. Before asserting that a model has clinical utility, it should undergo one final research step. This final step is an assessment of the model's impact on health outcome and potential cost-effectiveness.[218] Impact studies take an externally validated model and test the model prospectively in the real world. These studies can be challenging to do, time-consuming, and potentially expensive. Accordingly, very few models ever undergo a clinical impact study. In a review of 84 prognostic models published in six major journals between 2006-2009, 61 described development, 21 validation and only 2 provided evidence of an impact study.[205]

Despite an apparent lack of impact studies, there are hundreds of models in use in healthcare every day around the world. The next section will describe a selection of published models for decision support in major bleeding and coagulopathy.

1.6 Clinical decision support tools for haemorrhage and coagulopathy

Early initiation of a DCR strategy reduces delays and improves survival (see **1.4.3**). Many models have been created to assist with the early identification of patients that require large volume blood transfusion. There are relatively fewer examples of models to support the identification of patients with TIC. The following description of published models is included to serve as relevant benchmarks for comparison to the model analysed in this thesis.

1.6.1 Decision support for haemorrhage

Several systematic reviews have identified multiple models to support the identification of patients that require major haemorrhage protocol activation.[219] An extensive review, published in 2016, describes 36 models each with markedly different performance despite often sharing the same input variables.[220] The systolic blood pressure (SBP) was

the most common input variable (31/36 models) and results of blood or radiology investigation were used in just under half (15/36 models). Few of the models had been externally validated (11/36), and no impact studies were identified. Most of the models identified are intended for in-hospital use only.[220]

A selection of the most commonly cited models and models that highlight a particular area of interest to the research within this thesis are now discussed (**Table 3**).

1.6.1.1 The Assessment of Blood Consumption score: in-hospital model

The Assessment of Blood Consumption (ABC) score[221, 222] is a well-known and highly cited score endorsed by the American College of Surgeons', Massive Transfusion in Trauma Guidelines.[223]

The ABC model has been used to identify high-risk patients for inclusion into a large clinical trial.[224] It has been externally validated at least ten times.[99, 222, 225-232] The model was derived from expert consensus and is an example of a non-weighted score designed for rapid use at the bedside. The ABC has just four input variables, and the presence of each is worth one point.

A retrospective comparison performed in 2018, between the ABC score and clinician judgement, provides some insight into the score's potential impact. The study of 3421 patients demonstrated that the ABC score over triaged patients (higher number of false positives) compared to clinicians. For the use of ≥5 units of PRBCs in 24 hours, the clinicians had a positive predictive value of 65% vs the ABC score's 34%. Approximately half of the clinician MTP activations took place after the patient had left the ED. In these later MTP activations, the ED ABC score was positive in 81%. The authors infer that the use of the score would lead to a 35-minute reduction in MTP activation but extra blood product wastage.[233] The study does not report the true number of major haemorrhage patients, and so no sensitivity comparison is possible.

The score has been used to guide pre-hospital transfusion[152] and early warning to the receiving trauma centre.[234] An abstract that details the adaption of the ABC score for pre-hospital use is instructive as it highlights the necessity of a model to fit into the

clinical workflow. To operationalise the score pre-hospital a significant addition to the clinical practice of the flight nurses was required. The nurses were trained to conduct Focused Abdominal Sonography in Trauma (FAST) examinations during transport.[235] Without a FAST examination, the ABC score cannot be used as it is one of the four input variables. A pre-hospital ABC score was calculated using the same standard ABC score variables.

1.6.1.2 Trauma Associated Severe Haemorrhage Score

The Trauma Associated Severe Haemorrhage (TASH) Score[236] is one of the most cited massive transfusion prediction scores and has been externally validated seven times.[99, 225-228, 231, 237] It is a weighted-clinical score derived using a regression technique and is more complicated to calculate than many haemorrhage scores. It requires in hospital laboratory and imaging data. The score was developed and validated in a cohort of 6044 patients from the German Trauma registry entered between 1993-2003. Interestingly, the score was re-trained on a more contemporary set of patients to adjust the weighting of individual variables in 2011.[237] The TASH score is designed for use in the Emergency Department. The score often has very high discrimination when multiple models are compared.[226, 228] The score requires a blood gas and FAST examination. As such, it may not be able to output a risk score in a useful time frame. The TASH score is the only score known to perform better than clinicians in a prospective study of bleeding trauma patients (AUROC 0.72 vs 0.62; $p = 0.01$).[99]

1.6.1.3 Mina Score

A model devised by Mina, *et al.* in 2013 used a different approach to many of the scores published. This score uses the least absolute shrinkage and selection operator (LASSO) technique to develop the model. The LASSO technique combines variable selection with regularisation, i.e. weighting the variables. LASSO is commonly implemented in machine learning approaches. The resulting model has complex variable weighting, and the authors recommend the model is used within a mobile application ("an app") to perform the prediction.[238] Given 44 potential input variables, the LASSO technique selected four influential factors; MOI, HR, SBP and BD. The model does not require

dichotomisation of the input variables and so does not suffer from the loss of predictive power that dichotomisation causes.[239]

The model was developed and internally validated on retrospective data from one institution, and the model's output is the risk of activating the MTP. The model does not predict the volume of blood transfused or the amount of blood a patient requires. Instead, the model is trained to predict the decision to activate an MTP. An external validation was performed with the prospectively collected PROMMTT study patients.[240] In external validation the model's predictive accuracy was assessed in five different critical haemorrhage definitions. The model's discriminative performance dropped in this external validation from AUROCs 0.97 to ~0.70.[231] One reason for this decrease in performance is the PROMMTT study included only patients that had at least one unit of blood transfusion. However, the Mina score was trained using a dataset in which 96% of the population did not activate the MTP. Additionally, this external validation tested a clinically more useful outcome: the volume of blood transfusion a patient received. In combination, these two factors result in a different impression of the model's performance.

1.6.1.4 Shock Index

The Shock Index (SI) is the simplest of all models. It is merely the ratio of heart rate to systolic blood pressure. There is no consensus for the optimal threshold value to define a "positive SI" although >0.9 is the most commonly used.[241] In some populations SI outperforms the ABC score without the technical skill required for a FAST scan.[229] Four studies assess the utility of Shock Index to predict transfusion pre-hospital.[242-245] At a cut off of >1.0 SI is straightforward to calculate and has reasonable specificity (specificity 83%, sensitivity 57%).[244] This makes the test a useful rule of thumb to consider ruling out the need for massive transfusion but lacks the performance to aid decision making substantially.

1.6.1.5 Code Red

The "Code Red" score was developed in London to identify seriously injured patients who required transfusion on arrival at the hospital.[148] The score was developed using

expert knowledge, rather than data, and has three binary elements; suspicion or evidence of active bleeding, SBP <90 mmHg, and no haemodynamic response to a fluid bolus. The authors state that with the code red score, an experienced pre-hospital provider can accurately predict which patients will require blood products with a 91% positive predictive value (PPV).

Code red has been externally validated in Scotland, but similar to the original development study only code red positive patients are included in the analysis, so no specificity metrics are known.[246] The development and validation have further methodological issues that make the actual value of the score unknown. For example, there is no way of knowing whether the clinicians in the development study strictly applied the code red score to achieve the 91% PPV. In the Scottish validation study, a third of patients who triggered the code red were not hypotensive; one of the three scoring variables. It is not clear what additional diagnostic value is gained by the addition of two criteria above the clinical diagnosis of an expert. It is also unclear how generalisable the model's performance is when it is reliant on expert clinical diagnosis. The extent of framing bias in which the in-hospital transfusion of blood is influenced by the prior pre-hospital "code red" declarations is unknown. Pragmatically the model has become redundant as the practice of giving fluid boluses pre-hospital is no longer supported.

1.6.1.6 Red Flag

Red Flag[247] is a logistic regression derived score developed in France for bluntly injured patients. It is a bespoke pre-hospital non-weighted clinical score. The model's inputs include shock index, pre-hospital point of care haemoglobin measurement and pre-hospital intubation. The authors used a split sample development and validation approach, i.e. the external validation was a temporal validation from the same database. The score is limited it is application to blunt injury only and has no geographical external validation.

1.6.1.7 McLaughlin Score

Several models have been derived from the recent conflicts in Iraq and Afghanistan. Three models arose from US databases; McLaughlin[248], Larson[249] and Schreiber[250]. The MASH score[251] was developed from a British military cohort.

The McLaughlin is the highest cited of these scores. It is a non-weighted simple clinical score with four input variables; HR>105, SBP <110, pH <7.25, and haematocrit <32%. The model is trained only on patients that received blood. A bespoke military model reflects a need for models to be specifically developed in the patient population they are intended for. Differences in the mechanism of injury and patient demographics between civilian and military patients cause variability in model performance between patient populations.

1.6.1.8 MASH Score

The MASH score is the most recently developed score for military casualties.[251] It has been developed in an attempt to improve on the performance of other military models. The score is intended to "rule in" patients that need a massive blood transfusion, using only those variables available within the first few minutes of patient arrival to a medical treatment facility. The rule is designed to be clinically acceptable with input variables selected with regression methods after approval in a Delphi study.[252] The resultant model is an unweighted score with 13 input variables relating to injury, physiology and, unusually, previous treatment, e.g. pre-hospital blood transfusion. The score is temporally validated but not fully externally validated. In temporal validation, the score performed with high discrimination (AUROC = 0.93)

1.6.2 **Summary of decision support for haemorrhage**

Over 40 distinct models have been developed to assist decision making in major haemorrhage. The models vary in their methodological design, complexity, input variables, training population, development and validation sizes, intended location of use, the definition of the outcome and performance. The sheer volume of effort and different approaches is indicative of the need for accurate decision support in haemorrhage and the absence of agreement so far on the best model.

For pre-hospital decision support on haemorrhage several models cannot be applied as they require variables which are not available in the pre-hospital environment. For example, a score that depends on known laboratory values or radiology results does not produce useful results pre-hospital.

Despite the academic endeavour to develop and validate these models, there is no published prospective evaluation of their impact on clinical care. Therefore, their effect on decision making, resource use and patient outcome is unknown. Notably the NICE guidelines for major trauma concluded that there was insufficient evidence to support the use of *any* of these tools.[253]

All of these models share a challenging limitation as a result of the data available for model development. Haemorrhage decision support models trained on trauma registry or clinical trial datasets include the volume of blood the patient *received*. However, there is an important distinction to be made between the volume of blood a patient *received* and the volume of blood they *required*. It is far harder, but yet more clinically useful, to quantify how much blood a patient requires. As a result of training models from trauma registry data, the models predict an outcome of a clinical decision for a treatment, rather than the underlying pathological state. In other words, the models predict the volume of blood transfused rather than the volume of blood the patient has lost. As medical practice changes over time, a model which provides a guide on what treatment was previously given is less informative than a model that reflects what the patient requires.

The ideal model will be accurate, easy to use, and provide an early prediction. The model should not dichotomise variables and will therefore need to use modern computational

techniques. The model must be clinically acceptable. In this setting, a model which prioritises sensitivity (i.e. reducing the number of false-negative results) has greater utility than higher specificity (less false positives). The benefit of higher sensitivity is a result of the relative costs of the two false results. Most clinicians would argue that the risk of delaying or entirely missing a patient who required an early transfusion is higher than the financial and logistic costs associated with mobilising blood products unnecessarily. In an austere environment with limited resources, this benefit may carry more cost and be less desirable.

Table 3: Features of eight Major Haemorrhage decision support tools

	ABC	TASH	Mina	SI (PH)	Code Red	Red flag	McLaughlin	MASH
Publication Year	2009	2006	2013	2011	2016	2018	2008	2018
Intended Location	IH	IH	IH	PH	PH	PH	Military IH	Military IH
Internal validation (n)	596	6044	13961	n/a	126	3675	396	1186
External validation (n)*	5147	5147	1245	535	53	2999	596	No
Impact study performed	retrospective	No	No	No	No	No	No	No
AUROC in external validation	0.76	0.89	~0.70	0.77	unknown	0.83	0.85	unknown
Major haemorrhage definition	≥10 PRBCs in 24 hrs	≥10 PRBCs until ICU admission	#	≥10 PRBCs in 24 hrs	≥1 in hospital PRBC or haemorrhagic death	^	≥10 PRBCs in 24 hrs	≥6 PRBCs in 4 hours or ≥10 PRBCs in 24 hrs

AUROC; Area under the receiver operating curve, IH; In hospital, PH; Pre-hospital, PRBCs; packed red blood cells

* the largest or most representative external validation has been selected where multiple external validations have been performed. Temporal external validations were not considered true external validations and have not been included.

the external validation tests the model using 5 definitions of major haemorrhage; 10 PRBC units in 24 hours, Resuscitation Intensity score ≥ 4, critical administration threshold, 4 units PRBCs in 4 hours; and 6 units PRBCs in 6 hours

^ Multiple definitions used: transfusion in ED, or transfusion ≥ 4 PRBC in the first 6 h, or lactate ≥5 mmol/L, or immediate haemostatic surgery, or interventional radiology or death from haemorrhagic shock.

1.6.3 **Decision support for trauma induced coagulopathy**

In contrast to the numerous haemorrhage decision support models, coagulopathy prediction has received relatively less attention. To date, there is one in hospital and two pre-hospital prediction models for TIC. A further model discussed at 1.6.3.4 includes TIC but has been developed to predict patients in need of DCR.[254]

1.6.3.1 Cosgriff Score

The Cosgriff Score was the original coagulopathy prediction score.[255] It was developed in 1997 from a hospital registry of 58 patients. All the patients received >10units of PRBCs in 24 hours and had a GCS >8. Cosgriff classified patients as coagulopathic if they had twice the normal Prothrombin Time (PT) and Partial Prothrombin Time (PPT). With this classification of coagulopathy, 27/58 patients were coagulopathic. The model is a non-weighted clinical score derived using regression. The score's four input variables are pH < 7.10, temperature < 34ºC, injury severity score (ISS) > 25, and SBP < 70mmHg. The model was the first to demonstrate that coagulopathy may be predictable from clinical information. (**Table 4)**

The model has flaws which reduce its ability to support real time decision making. One of these flaws is the inclusion of ISS as an input variable. The ISS can only be calculated once all the patient's precise anatomical injuries are known.[256] The ISS has 2000 coded injuries and is typically calculated by trained research staff with the aid of the scoring manual, days after the patient's admission. The authors speculate that clinicians will be able to estimate the ISS. However, the ISS is not a score that clinicians regularly use. As a result, the accuracy of clinical estimation is unknown. Regardless of the practicality of calculating the score, the model did not separate the patients into two clinically distinct risk groups. Regardless of their coagulopathy prediction included patients had high rates of transfusion and mortality. The non-coagulopathic group had a mortality rate of 42%. This lack of clinical risk stratification is due to the patient selection criteria. These fundamental flaws in the model's development result in a model that cannot be used to support decisions and has not been incorporated into clinical practice. The model has not been externally validated.

1.6.3.2 The Coagulopathy of Severe Trauma (COAST) score

In 2011 Mitra, et al. published the coagulopathy of severe trauma (COAST) score.[257] In 2019 the COAST score was incorporated into practice and used by the Queensland Ambulance Service.[258] the COAST score is used to support decisions about giving TXA and identifying patients for the PATCH trial (NCT02187120).[259]

The COAST score offers many advantages over the previous Cosgriff score. COAST is derived from a large representative sample of trauma patients and is designed to identify patients with coagulopathy using variables available pre-hospital. COAST is a simple weighted score developed using regression. The development of the score did not exclude patients with known causes of coagulopathy. For example, patients on oral anticoagulation or liver disease were not excluded. The model has two continuous variables (SBP and Temperature) which are dichotomised in the COAST score. The other input variables are patient entrapment, chest decompression and abdominal or pelvic content injury. The maximum score is seven.

The pre-hospital diagnosis of abdominal or pelvic content injury is a clinical diagnosis and therefore requires specific attention as it is prone to error. The authors report a 75% sensitivity and 95% specificity for the pre-hospital staff's diagnosis of injury. This re-enforces that there is a group of patients with occult abdominal and pelvic injuries. These patients with occult injury, are at risk of being underscored by the COAST score.

A COAST score >2 is classified as positive. For the prediction of coagulopathy in the internal validation cohort a threshold >2 results in a sensitivity of 70% and specificity of 84%. The 40/100 patients that were false negatives had lower ISS, less need for urgent surgery, lower mortality and a higher proportion of isolated severe head injury. Nevertheless, these false negative patients still received large volumes of blood transfusion and had a mortality of 28%. While the false negative patients appear to have a different and potentially less severe injury pattern, they are still at high risk of death. A model that appears to systematically underscore less severely injured patients might have poor calibration. Unfortunately, the authors do not report calibration and so full evaluation of the model is not possible.

The authors published a small external validation in 2019 using 133 patients from a trauma centre in Belgium.[260] The Belgian dataset did not have exactly the same variables as the development dataset, and so the validation uses one different input variable: chest decompression was replaced by major chest injury as suspected or diagnosed pre-hospital. Unusually, in external validation, the model performed with better sensitivity and specificity than in development. A larger external validation ought to be performed before the performance of the model is demonstrated to be generalisable. Nevertheless, the study hints towards the ability to separate groups of pre-hospital trauma patients using a simple clinical score. The clinical impact remains unknown.

1.6.3.3 The Prediction of Acute Coagulopathy of Trauma (PACT) Score

Most recently Peltan *et al.* published the Prediction of Acute Coagulopathy of Trauma (PACT) score.[261] The authors sought to improve on the predictive accuracy of the COAST score. The PACT model is a more complex weighted clinical score derived using logistic regression. Their approach to develop and validate the model is methodologically strong. The publication conforms to the 2015 TRIPOD guidelines which facilitates transparent critique of the model.[211]

Some of the methodological strengths of the PACT model include i) using a mixture of data-driven and clinical knowledge methods to select predictor variables, ii) multiple imputation of missing variables, iii) use of a majority rules algorithm to balance overfitting with predictor variable selection, iv) mitigation of overfitting by using an assessment of model optimism with bootstrapping, v) sensitivity analysis to classify patients with TIC, vi) a large development dataset and vii) an independent external validation cohort.

The authors developed the model using data from 1963 severely injured patients in the Oregon Trauma Registry. An external validation was conducted using 285 patients. These patients were part of another prospective study and all attended the Harborview Level 1 trauma centre. Criteria for inclusion into the validation cohort were similar to the development cohort. Patients in the validation cohort were more severely injured by

ISS and were included if they had a blunt injury and received at least a unit of PRBCs within 24 hours of injury.

The PACT model (www.pactscore.com) input variables are age, pre-hospital cardiopulmonary resuscitation (CPR), intubation, pre-hospital GCS, shock index, and non-vehicular injury mechanism. The model's performance was sustained in the external validation cohort and appeared to perform better than the COAST score. The authors set an operating threshold which optimised both sensitivity and specificity. At the selected cut-off the model identified 97% of patients without coagulopathy and 22% of those patients with coagulopathy.

The PACT model has been well developed using appropriate statistical techniques. Importantly, the authors ensured they developed their model in the type of population for its intended use. The patient population is a severely injured cohort rather than an unselected trauma population. In the unselected trauma population, the vast majority of the patients will have very low risk of coagulopathy and the discrimination of a model may appear very high.

1.6.3.4 TICCS

The Trauma Induced Coagulopathy Clinical Score (TICCS)[262] was developed as a simple clinical score for paramedics to use on scene. The TICCS score predicts risk of requiring DCR. To classify which patients require DCR, the authors used a range clinical states; acute traumatic coagulopathy, haemorrhagic shock, massive transfusion or requirement for a haemorrhage control intervention. The input variables were selected based on knowledge rather than being statistically driven. In contrast to any other TIC score, TICCS provides an input variable that accounts for the clinical impression of the patient's injury severity by the treating clinicians. It is not clear how the weights of the input variables were selected.

Development was performed in a prospective study using data from 82 patients. Inclusion criteria into the development population was "severe trauma" and is not further expanded. The model has been externally validated twice but only to assess the

performance of the model for the prediction of major haemorrhage.[263] [264] The ability of this model to identify patients that require DCR or have TIC, outside of the development cohort, is unknown.

Table 4: Features of four acute traumatic coagulopathy decision support models

	Cosgriff	COAST	PACT	TICC
Publication Year	1997	2011	2016	2014
Intended Location	IH	PH	PH	PH
Development Cohort description	GCS>8 & received > 10 PRBC in 24hrs 1993-1995 1 x Level 1 Trauma centre, USA	ISS>15, or immediate surgery, ICU or death 2006-2008 1 x Level 1 Trauma centre, Australia	severely injured* Excluded iTBI 2008-2012 Multicentre (44) USA	severely injured 2012-2013 Single Centre Belgium
Development (n)	58	1680	1963	82
Internal validation (n)	n/a	1225	Bootstrap from development	n/a
External validation (n)	n/a	133	285	no
Impact study performed	no	no	no	no
AUROC in external validation	n/a	0.94 (CI: 0.88 - 0.99)	0.80 (CI 0.72 - 0.88)	no
TIC definition	PT AND PPT >2x normal	INR > 1.5 OR aPTT >60	INR > 1.5 on 1st ED sample	Not only TIC^

GCS; Glasgow Coma Scale, PRBC; Packed red blood cells, PT; Prothrombin Time, PPT; Partial Prothrombin Time, INR; International Normalised Ratio, ICU; Intensive care unit, ISS; Injury severity score, iTBI; isolated Traumatic Brain Injury.
* Severe injury was defined as death prior to discharge; admission directly to the ICU or operating room from ED; or transfer from the initial ED to another state-certified trauma centre ICU or operating room.
^ Model trained to identify patients for DCR. Coagulopathy was one criterion for DCR. TIC classified using ROTEM, INR (>1.3) and fibrinogen (<1.5 gL^{-1}) measurements. ROTEM measurements to classify TIC were ≥20% derangements in ROTEM ExTEM clotting time, clot formation time, maximum clot firmness and maximum lysis.

1.6.3.5 Individual parameters to predict coagulopathy and transfusion pre-hospital

A noteworthy analysis was conducted by David *et al*.[265] The purpose of this study was to assess the individual relationship of a given variable to one of two outcomes; TIC and massive transfusion. In a retrospective cohort of 485 trauma patients from a French trauma system the authors identify the most discriminatory variables to predict coagulopathy. These variables were ISS (AUROC: 0.84), pre-hospital fluid volume >1000ml (AUROC: 0.80), Vasopressor administration (0.75) and shock index (AUROC: 0.72). This study highlights similar predictor variables included in other models and reflects similar findings from a distinct trauma system.

1.6.4 **Summary of decision support of trauma induced coagulopathy**

Prediction of trauma induced coagulopathy has been a research goal for over 20 years. Over that time the methods used to develop the models have become more rigorous and model performance has improved. The intended use of TIC prediction models has shifted towards improving patient selection into clinical trials and away from identifying which patients should undergo TIC treatment.[266, 267]

There are several possible reasons for the move towards stratifying patients for research rather than treatment, the foremost of which is a lack of predictive accuracy of the current models.[266] Part of the reason for suboptimal performance is likely due to the reliance on regression models and conversion into simple clinical scores. Forcing continuous variables into binary options reduces predictive power by up to a third.[268] Statisticians responsible for the TRIPOD statement, agree that *"categorising continuous predictors produces models with poor predictive performance and poor clinical usefulness*."[268] They make the further important point that not only is categorisation statistically inefficient, but it is unnecessary and biologically implausible.[268] All the models to predict TIC categorise their continuous variables.

There are other methodological concerns with regression derived scores which impact their performance. None of the models used non-linear relationships (such as fractional polynomials or restricted cubic splines) or evaluated interaction terms between

variables. It is unlikely that all the relationships are truly linear and unrelated to each other. Another assumption of regression, which these models violate, is that the variables are normally distributed in the population. As a result of these assumptions and violations the models lose predictive power. On a practical level, simple clinical scores cannot produce meaningful risk stratification if an input is unknown. For example, a GCS score without the voice component is not a useful metric. Imputation methods are possible, but unlikely to provide a solution to missing data in a simple clinical score when used prospectively.[269]

Finally, the models are ostensibly designed to predict TIC. However, they have been developed and validated to predict the result of a coagulation test. INR, PT and aPPT/PPT are measurements of clot formation time. They are incomplete measurements and do not assess biologically important processes such as fibrinolysis. As such there are multiple threshold values of these laboratory tests used to classify coagulopathy (see **7.5.4**). Using these measures does not identify the true state of the coagulation system. Models developed to predict these test results compound measurement error and reduce the clinical utility of these models. It is not the INR that is important, rather it is knowing the true underlying state of the coagulation system that provides an opportunity for intervention and improved patient outcome.

1.6.5 **The lack of impact studies for decision support**

Sections **1.6.1** and **1.6.3** detail the development and validation of 12 decision support models for haemorrhage and TIC. Together these publications have thousands of citations yet not a single model has undergone a comprehensive impact study. An impact study quantifies the effect of a model on decision making and patient outcomes in clinical practice using a comparative design.[218, 270] Without an impact study it is impossible to know what the effects of a model will be on patient or system outcomes.[271, 272] There are several aspects to consider in an impact analysis: i) was the performance of the prediction rule preserved in clinical practice? ii) was the realised impact more or less than the potential impact (recommended decisions compared to actual decisions)? and iii) overall did the model impact patient care?

The gold standard for impact studies is to cluster randomise clinicians into a group with the prediction model and compare to a standard care group. Such trials are often expensive and time consuming. However, when they are conducted, they can uncover unexpected effects and provide necessary information about how to implement decision support successfully.

Well conducted impact studies produce useful information. For example an impact study of nearly 800 surgeons in the USA demonstrated an unexpected result.[273] The study intervention was a model to predict operative risk. Surgeons were randomised into two groups; with and without the model. In the study the surgeons were given four clinical vignettes and asked to give their likelihood of recommending an operation. The study demonstrated a difference in the perception of the risk associated with an operation between the groups. Despite this difference the rate of proposed surgical intervention was the same between the two groups. The authors inferred that while the risk calculator only gave information relating to the risk of surgery, the surgeons were making a parallel adjustment of their perception of *non*-operative risk. This unintended and off target change had the effect of neutralising the impact of the model in the decision for surgery.

Impact studies have also demonstrated that the strategy used to communicate the advice of a CDST influences the clinical effect. This point is illustrated by two prospective impact studies that examined the use of a post-operative nausea and vomiting (PONV) risk score. In the first study, the risk score was given as a percentage. In this study the rate of anti-emetic prescriptions increased but the rate of PONV did not change.[274] In the second study the risk score was given to clinicians as a treatment recommendation. In this study the rate of PONV decreased in the intervention group.[275] Together these studies suggest a directive approach (instructing the clinician to administer a treatment) may have a greater impact on patient outcome than the theoretically more satisfying assistive approach.

It is clear that the introduction of decision support into clinical practice is a complex intervention. Impact studies can be expensive and time consuming but provide necessary information to assess the clinical utility of decision support. Without an impact study the clinical effect of a decision support tool will be little more than a guess.

1.7 Machine learning in medicine

1.7.1 The rise of machine learning

In 1965 Gordon Moore, a co-founder of Intel, predicted that every two years the number of transistors on a microchip would double and the price of computers will be halved.[276] His prediction, now realised as Moore's Law, underpins the dramatic increases in speed and accessibility of computational power. The dramatic advances in computing performance effect nearly every element of modern society; from art to justice, transportation to communication and business to medicine.[277] [278]

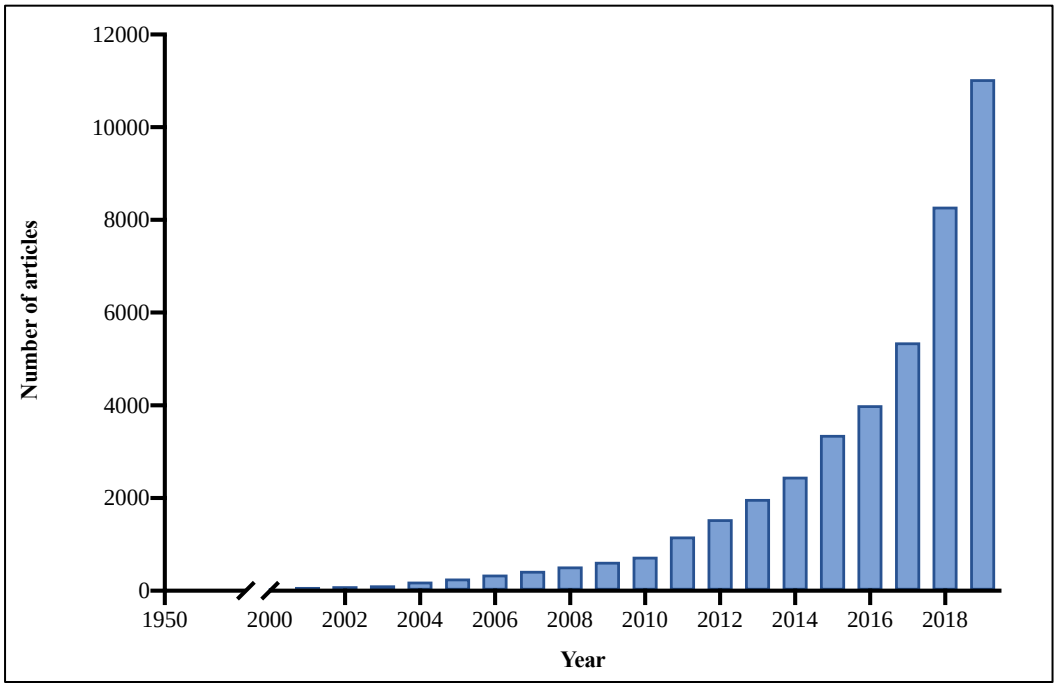

Figure 8: Histogram of the number of articles in the US national Library of Medicine, National Centre for Biotechnology Information (Pubmed.gov) by year. Articles were identified with the search term "Machine Learning". The search was performed 8th April 2020

The 2017 the British Government commissioned an external review to establish the likely impact of digital technology on the NHS in the next 20 years. The report was published in 2019 and specifically addressed advances in genomics, digital medicine, robotics and artificial intelligence.[279] Known as the Topol review after the senior author,

the review concludes that medicine is in a unique position with massive data being generated through electronic patient records, medical imaging, biosensors and smartphone applications. It is these large datasets that provide the substrate for artificial intelligence to make vital inferences. This combination of sensing hardware, digitally readable information and powerful computation will pave the way for more rationale, efficient and personalised healthcare.

The Topol report defines Artificial Intelligence (AI) as the science of *"getting computers to do tasks that would normally require human intelligence"*.[279] The concept of AI is not new with its roots in the 1950s (**Figure 8**).[280] A key element of Artificial Intelligence is machine learning. The Topol report borrows a definition of machine learning (ML) from a Royal Society publication: *"Machine learning is a branch of artificial intelligence that allows computer systems to learn directly from examples, data and experience"*.[281]

The boundaries between machine learning and statistics are blurred. The approaches might be better view on a spectrum. ML relies heavily on computational power in a way that statistics does not. Traditional statistics relies on small samples and strong assumptions about the data. In the purist forms, machine learning is able to learn from data in ways that traditional statistical approaches cannot. Many statistical techniques such as regression modelling, require assumptions about how variables are distributed and related to one another. For example, regression methods assume input variables are normally distributed, independent and have equal variance. When these assumptions are violated, the resultant models have reduced predictive performance. Machine learning methods are not constrained in the same manner and there is no need to restrict the number of variables. Interactions and non-linear relationships between variables can be easily identified with machine learning, in ways that are prone to human bias with traditional statistical techniques.[282]

1.7.2 **Algorithms: the language of machine learning**

Machine learning is driven by algorithms. Algorithms are the set of rules a computer follows to solve a problem. As well as increased data and computational power, the performance of algorithms has improved in the last 10 years. With increased ability to learn from data, machines have surpassed human ability in some applications. **Figure 9** demonstrates the improvement of machine classification of images over time. By 2016 machine learning had exceeded human image classification.

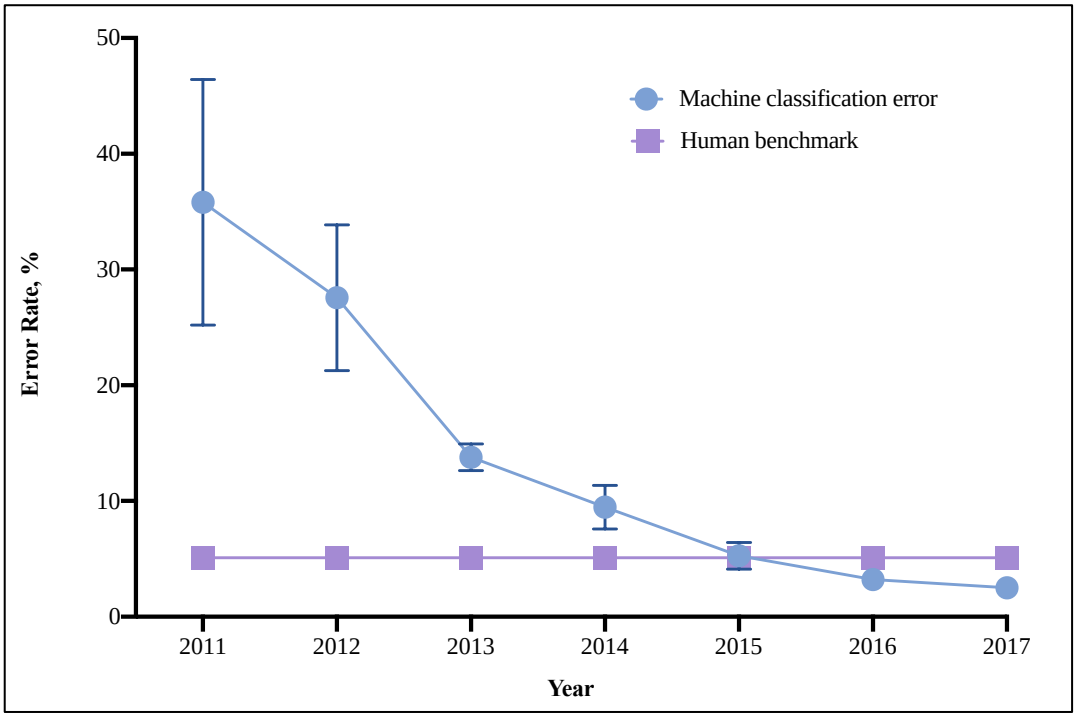

Figure 9: Results of the ImageNet challenge, error rate of image classification by year. Blue line denotes machine performance. Magenta line denotes human performance.[283]

There are many distinct algorithms used for machine learning. This diversity is representative of the wide range of data types, structures and relationships which require different analytical approaches. Broadly, algorithms can be grouped in two ways: their learning style and by similarity of form and function. An algorithm's learning style can be supervised, unsupervised or through reinforcement. Supervised learning requires the data to be labelled by a human and is the most common approach used in medicine. The machine learning community consider regression to be an example of a supervised

learning technique. Supervised learning algorithms build a mathematical model from training data containing both the required inputs and outputs. Once the model has been constructed, it can be used to determine the unknown outputs of new input data.

Unsupervised learning does not have human labels on the training data. Instead the algorithms learn statistical relationships between input data and subsequently label similar data points. Clustering is a common example of unsupervised learning. The third learning style is reinforcement learning, in which, algorithms are trained to optimise a reward outcome. For example an algorithm trained to play "Breakout", an Atari computer game, was trained to optimise its score by hitting bricks with a ball. Notably the algorithm quickly exceeded the performance of the human expert.[284]

Within the different broad categories of algorithms (supervised, unsupervised and reinforcement) there are many different algorithms, each with their own strengths and weaknesses (**Figure 10**). Deciding which algorithm to use depends in part on the size, quality and nature of the data and also how the results of the algorithm will be implemented.[285]

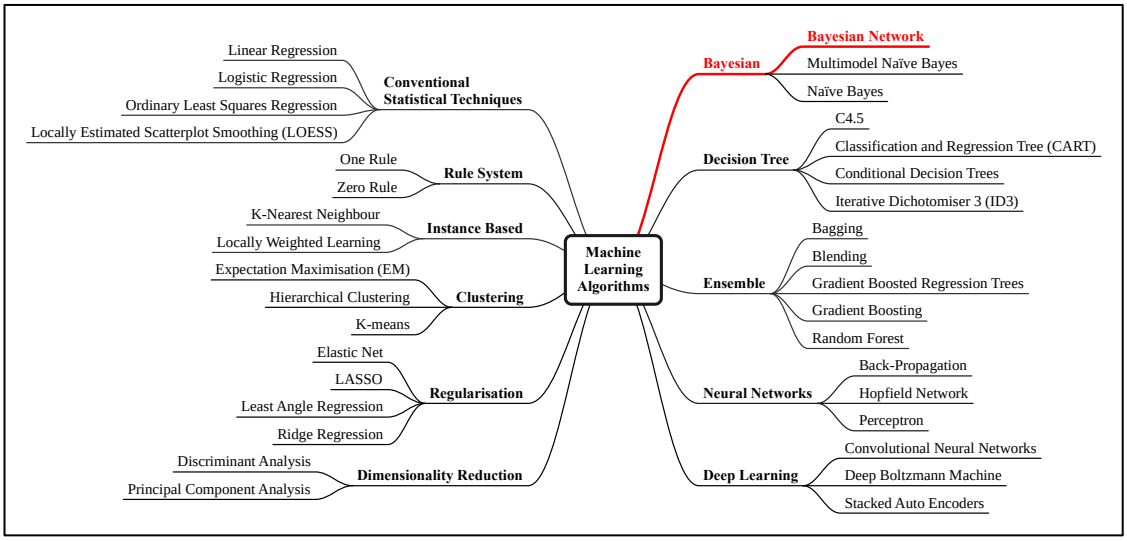

Figure 10: A mind map of algorithms grouped by similarity Bayesian Networks are highlighted in red as the central machine learning method used in this thesis.

1.7.3 **Application of machine learning to medicine**

The most celebrated examples of the application of machine learning in medicine relate to the use of deep neural networks (DNN) in image classification. Radiology[286, 287], pathology[288], fundoscopy[289, 290], dermatology[291, 292], endoscopy[293] and echocardiogram interpretation[294] all produce images that can be interpreted with machine learning. In recent years programmers have started to compare their algorithms with the performance of doctors.

The comparison of man versus machine in medicine is often overly reductionist on the part of the human. For example, a comparison of four radiologists to a 121-layer convolutional neural network in detecting pneumonia from chest x-rays put the performance of the machine ahead of the clinicians.[295] An important caveat is that radiologists do not simply assess a chest X-ray for one diagnosis. Neither do they perform interpretation of radiological images in a vacuum of patient information. Furthermore, while the neural network performed well, the discriminatory ability was not high enough for the machine to be employed autonomously (AUROC = 0.76).

Another approach, rather than man versus machine, is the assistive model of human and machine integration. This type of study was used recently to assess the impact of a DNN for fracture identification on X-rays. Using a DNN trained in part by information from orthopaedic specialists, emergency medicine clinicians improved their diagnostic accuracy of wrist fracture detection from 81% to 92%.[296]

Machine learning models that do not display perfect accuracy can still be useful. A randomised controlled trial (RCT) in a simulated setting demonstrated a deep learning algorithm was able to "read" an acute CT head scan 150 times faster than the radiologists (1.2 seconds compared to 177 seconds). If speed is the chosen performance metric, then computers will always outstrip human performance. The algorithm's diagnostic accuracy was measurably inferior but may still have a role in triage or screening for subsequent human assessment.[297]

The Topol report recognised that the application of machine learning to medicine will change the way medicine is practiced in the next 20 years. As these examples have demonstrated there are several likely relationships for humans and machines. For tasks that computers can do rapidly and efficiently machines will *replace* humans. In tasks which are predominately straight forward for machine learning but have occasionally difficult cases, machines may *filter* for human assessment. In very complex or high risk situations that require human judgement computers will *augment* human activities.

1.7.4 **Application of machine learning to haemorrhage and coagulopathy**

A review of machine learning in predicting trauma outcomes was undertaken by authors at the US Army Combat Casualty Care research programme in 2017.[298] The review identified 65 studies of which 12 were on the prediction of morbidity, shock or haemorrhage. There were no applications of ML to the prediction of TIC. For the prediction of haemorrhage the studies included an ANN for the prediction of transfusion in the ED[299], analysis of continuous vital signs[300 301] and machine enhanced analysis of diagnostic tests such as plethysmography[118, 302-304], FAST[305] and doppler.[306]

Since the review was published, a study employing an ensemble algorithm called SuperLearner was applied to a cohort of trauma patients enrolled in an observational study. The algorithm identifies and autonomously selects the key input variables from the dataset. SuperLearner builds models to optimise prediction for a given output using different autonomously selected algorithms. Amongst the numerous outcomes this study describes, the prediction of coagulopathic trajectory was the weakest reported (AUC 0.48-0.88).[307]

The Compensatory Reserve Index (CRI) is a device built on a machine learning application and has progressed well along the developmental pathway (**Figure 7**).[118] The CRI uses a machine learning technique called feature selection to quantify changes in arterial waveforms using non-invasive photoplethysmography. CRI's feature selection assesses over 200 data points from one arterial waveform. Comparing the shape of the waveform to training data of simulated shock, the CRI is able to produce a quantification of the degree of compensation to shock. The CRI measuring device is

little more than an oxygen saturation probe. The authors have demonstrated the CRI to be more sensitive to blood loss then traditional vital signs.[303, 308] In a prospective observational study of 89 trauma patients at a single level 1 trauma centre, the CRI was more accurate in the prediction of haemorrhage than SBP (AUROC 0.81 and 0.62 respectively).[107] Further work is required to see if this performance is can be generalised to other populations and what the impact of the device is on medical care.

1.7.5 Limitations of machine learning for haemorrhage and coagulopathy

Many of the successful applications of machine learning both in health care and other wider fields use deep learning approaches.[309] Deep learning is best performed when the training data is very large, so called "big data". Big data refers to data that is collected often without a strict experimental design. Big data has proliferated in the last 5 years due to the ease of acquisition, storage and interconnectivity between devices.[281, 310] The storage of hundreds of thousands of images, each with vast numbers of pixels, has enabled powerful deep learning approaches to assist image recognition.

The volume of data from trauma patients to develop machine learning tools does not compare; it is substantially smaller. Novel methods of machine reading electronic health records are proposed. In the future such approaches may generate large volumes of data for trauma patients.[311] However, at present these methods are in their infancy. Datasets for developing models to assess coagulopathy and bleeding rely on highly structured, and expensive recording methods such as cohort studies and clinical trials.

Another important consideration for the application of machine learning approaches to medical practice is a balance between interpretability and accuracy of prediction. As machine learning techniques become more complex, the ability to interpret how the model arrived at a given output is increasingly opaque.[312] Healthcare providers unable to understand how a prediction has been generated are required to place their trust in the prediction. High profile machine learning failures, such as IBM Watson for oncology, erode trust in nascent technology.[313] Similarly, in a market with substantial financial rewards, doctors are quick to point out where machine learning approaches are failing or unsafe.[314, 315]

Systematic errors of machine prediction have a greater potential for harm than a single doctor's error. Whether rigorous regulatory processes will be enough to gain medical professional's trust in the technology remains to be seen. Governments around the world are devoting resources into improving the rigor of increasingly complex medical device regulation. [279]

Just like the traditionally derived decision support tools discussed in 1.5.3, machine learning approaches are not able to demonstrate their clinical utility simply by demonstrating high predictive performance with AUROC values. Problems with bias in training datasets limiting generalisability, for example, are well documented in many fields of machine learning.[277] Algorithms must undergo clinical impact studies in precisely the same fashion as traditional clinical decision support tools before their clinical worth can be determined.[316, 317] Examples of impact studies of machine learning in healthcare do exist[290, 293, 318] but are far from common place.

1.8 An introduction to Bayesian Networks

1.8.1 Bayesian philosophy and Bayes Theorem

Bayesian statistics are a fundamentally different approach to the well-known frequentist statistical methods advanced by Fisher, Pearson and Neymar.[319] At its core, Bayesian philosophy believes uncertainty must be described by quantification using probability. The degree of belief of a rational agent can be used to measure the likelihood of an uncertain event.[320] This is a radically different concept of uncertainty compared to frequentist philosophy. Frequentists see the probability of an unknown event as random and its likelihood can be determined by measuring its frequency in a set of experiments.[321]

Bayesian statistics precisely apply mathematical reasoning to uncertain situations. Bayes theorem incorporates the prior belief of the probability of an event and calculates how that prior belief should be updated to account for new evidence.[322] This process is known as Bayesian Inference. Initial beliefs are known as the prior probability and the

updated belief is the posterior probability. These conditional probabilities are expressed in Bayes Theorem:

This Theorem relates the probability of variable *A* given (or conditioned on) a variable *B.* In trauma this theorem could be used to represent the probability of a patient bleeding given a low systolic blood pressure measurement. Where there are few inputs to a problem Bayes theorem is straightforward to calculate. However, when there are complex problems with multiple influential variables the calculations become more elaborate and time-consuming to perform.

1.8.2 **Bayesian Networks**

Bayesian Networks (BN) are a machine learning application of Bayes Theorem. BNs are *causal probabilistic* models. A *causal* model describes the causal mechanisms of a system. Causal modelling provides thorough insight into risk by modelling cause and effect relationships. Causal models prevail over over-simplistic statistical approaches which neglect key influential variables.[320] A *probabilistic* model quantifies the uncertainty in the prediction as a probability. Probabilistic models are distinct from classification models.

The purpose of a BN is to determine the likely values of an unknown variable. For example, whether a patient is likely to be bleeding given four vital sign observations and their mechanism of injury. This calculation is performed with Bayes' Theorem and executed using commercial software packages such as Agena Risk.[323]

As described in **1.7.1** advances in computational methods provide the opportunity to develop complex models for outcomes with multiple influential variables. BNs provide a framework for combining multiple sources of information, including existing knowledge and individual patient characteristics, to accurately compute individualised risk estimates and even the likely effect of a therapeutic intervention.

1.8.2.1 Components

BNs consist of two parts (**Figure 11**). The first is a network structure which graphically describes the model's causal relationships. The network is graphed using a directed acyclic graph (DAG) in which nodes represent uncertain variables.[324] Arcs link the variables and represent probabilistic dependencies between variables. The second component of a BN is a set of parameters that describes the strength of the influences between variables. These parameters are known as the conditional probability distributions.[320]

BNs are flexible and both the structure of the DAG and the parameters that link the variables can be constructed in different ways. BNs can be developed purely from the data or "by-hand" (using experts to elicit knowledge) or via a combination of these two methods. Variables can be both continuous or discrete and may be directly measurable or unmeasurable.

BN are developed using reasoning pathways consistent with clinical understanding. This permits difficult to characterize physiological states to be incorporated, such as hypoperfusion. Portrayal of these *latent* variables allows a coherent arrangement of the directly observed values. This gives the end-user information about how the prediction is calculated and increases the model's generalisability to groups of patients beyond the training dataset.

1.8.2.2 Strengths of Bayesian Networks

Bayesian networks have advantages over more commonly used clinical decision support tools such as regression based models.[320] One of the most significant advantages of a Bayesian Network is the ability to combine domain expertise with data. Using experts, prior knowledge can be incorporated into BNs either by their influence over the network design and causal relationships or by assigning probabilities to relationships using published literature and their experience. Combing expert knowledge and data, produces accurate models.[325]

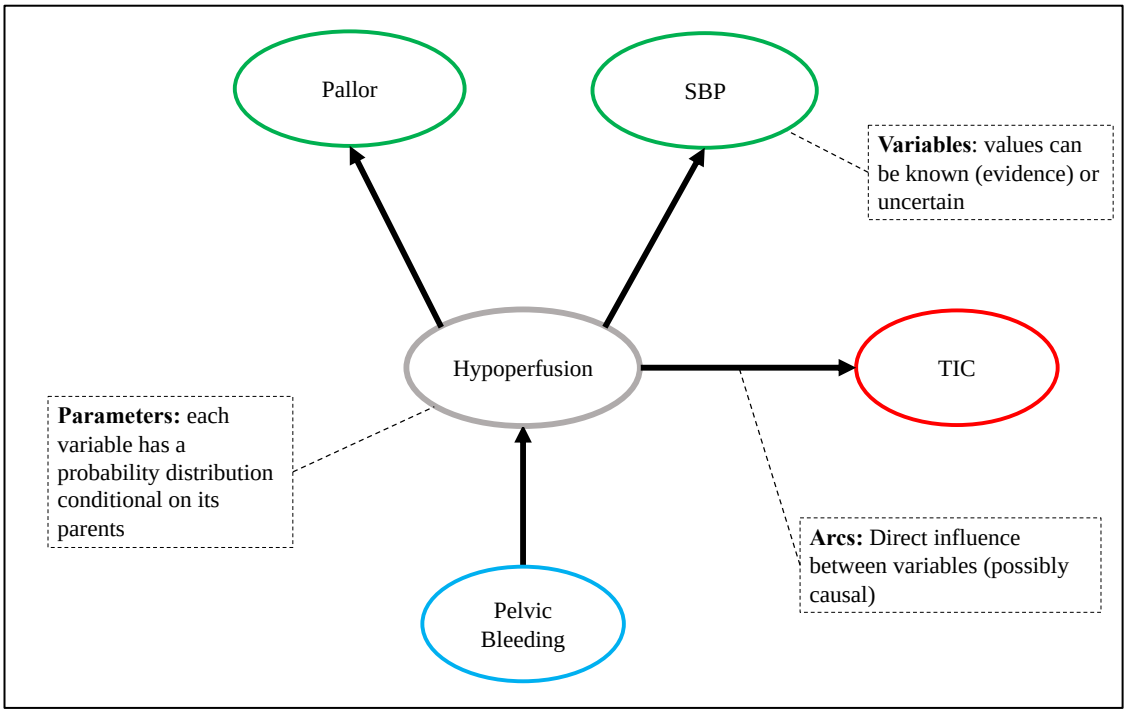

Figure 11: A simplified fragment of a Bayesian Network. The green nodes denote measurable indicators (evidence) of hypoperfusion. The grey hypoperfusion node is an example of an indirectly measurable (latent) node. The red node is the outcome of interest. The blue variable is a parent of the grey variable. Arrows demonstrate the direction of the causal relationships.

As the combination of knowledge and data can be simply represented as the BN's graphical structure, it is easy to explain and interpret the model's assumptions. With a transparent structure, clinical credibility can be established. It is possible that this transparency may subsequently enhance clinician trust and uptake of the model into clinical practice.

BN are able to model complex problems where a multitude of related variables interact, often in nonlinear relationships. BN models are less susceptible to violations of statistical assumptions common with logistic regression models. For example, in a BN model collinear variables do not produce the same degree of model instability seen with collinearity in regression models.

The causal structure of BNs also permits causal inferences to be made in a way that is not possible with other machine learning techniques.[326] This allows BNs to model causal interventions, and to reason both diagnostically and prognostically.[327]

In pre-hospital care, decision making under uncertainty due to missing or unknown information is common. With BNs, any subset of the input variables can be used for evidence. The BN will update the probability of all unknown variables when information from any variable is entered.

Finally, Dynamic Bayesian Networks provide the ability to make models that reflect decision making that is not static. For example, a clinician making a prediction of a disease may change their impression of the likelihood of that disease overtime as increasing information becomes available. In the context of trauma, a familiar example would be a pre-hospital physician who recognises a patient with low blood pressure after a fall from height. She may consider the cause of shock to be either blood loss or a spinal injury. Over time if the patient's observations become increasingly tachycardiac and hypotensive the clinician is more likely to believe the patient is bleeding and hence the probability of hypovolaemia as the cause of shock is increases. In hospital, following a positive FAST scan, the clinical team believe the patient has intra-abdominal bleeding with a high probability and rapidly move the patient for an urgent laparotomy. This example illustrates the evolving nature of medical decision making over time. Dynamic BNs can model this series of time dependent decisions in a way static prediction models cannot.

1.8.2.3 Limitations of Bayesian Networks

BNs are mathematically complex and their development is time-consuming compared to purely data-driven techniques such as neural networks. Often the most time-consuming element of building a Bayesian Network is determining the graphical structure and key variables for a given problem. This is because BNs are causal models and the relationship of the variables must be arranged in such a way that the causal relationships are respected. Some problems have very clear model structures whereas for other problems it may not be clear which variables represent cause or effect. Furthermore, it may not be clear whether a particular variable is involved in a causal relationship at all.[328] Hence, defining the network structure can be difficult and is subject to potential bias from the domain experts.

It may also be difficult for domain experts to express their knowledge in format that is compatible with a probability distribution. This may be due to inexperience with Bayesian methodology. Additionally, the task of estimating the probability of rare events is difficult.[328]

1.8.2.4 Examples of Bayesian Networks in medicine

Bayesian Networks have been developed to aid decision making in medicine for over 20 years.[329] They have been applied to problems of diagnosis, prognosis and treatment. Kyrimi *et al.* identified 47 examples of medical BNs[327], ranging from anticoagulant prescription decisions[330] to the diagnosis of ventilator associated pneumonia.[331] The most highly used applications of Bayesian networks are in identifying gene regulatory networks from time course microarray data[332], and the identification of breast cancer by combining patient risk factors with mammograms.[333]

Reflecting a similar situation to the lack of impact studies discussed in section **1.6.5**, Kyrimi's review did not identify any studies describing prospective external validation or impact studies for medical Bayesian Networks.[334]

1.9 The Perkins-Yet "ED TIC BN"

1.9.1 ED TIC BN model development

Perkins and Yet developed a Bayesian Network to predict TIC within the first 10 minutes of hospital care.[335-338] The clinical decision support tool is designed to be used in the Emergency Department (ED). The BN requires routine clinical information that is readily gained during a standard Advanced Trauma Life Support (ATLS) patient assessment process. The model's casual structure was developed using the "by hand" method. It is derived from evidence in published literature (**Figure 12**). The conditional probability distributions were trained using patient data from 600 patients in a civilian international, multicentre, prospective cohort study called the Activation of Coagulation and Inflammation in Trauma (ACIT).[339]

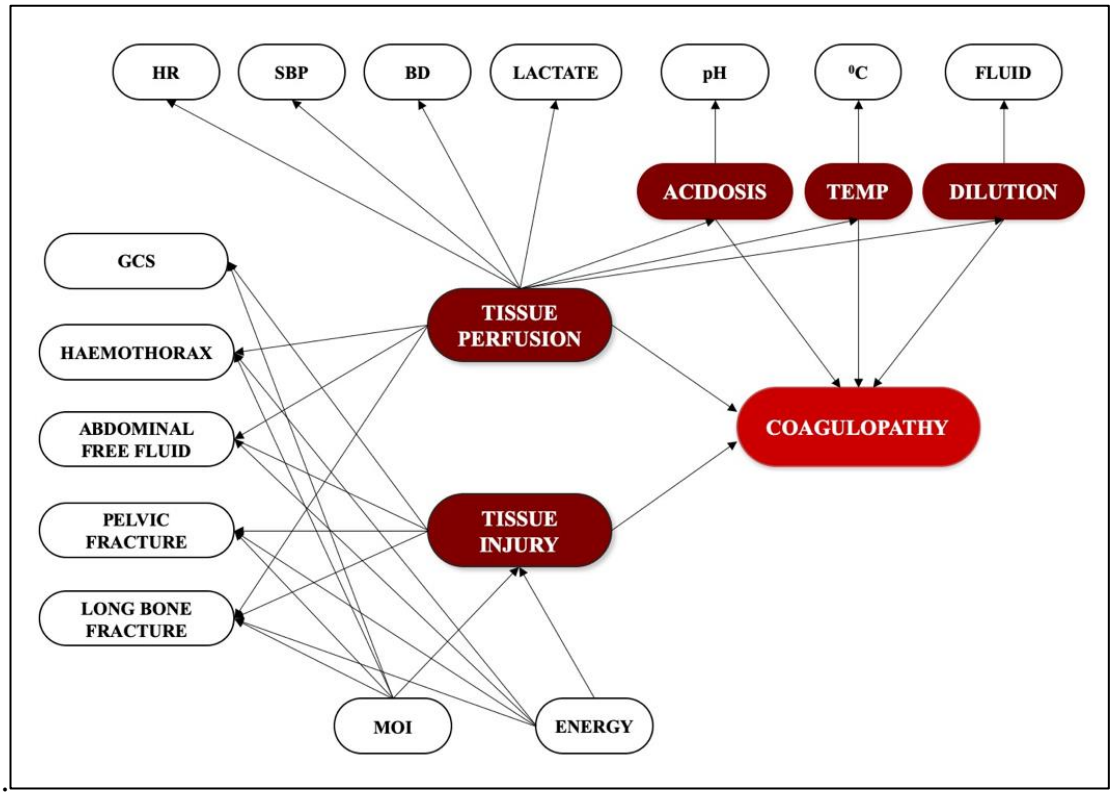

Figure 12: Directed acyclic graph of the Perkins-Yet Bayesian Network for prediction of trauma induced coagulopathy in the emergency department. Five latent variables (dark red nodes) are causally related to the outcome variable coagulopathy. The model uses 14 evidence measurements (white nodes) to compute the likelihood of coagulopathy. Fluid, volume of prehospital resuscitation crystalloid administered; GCS, Glasgow Coma Scale; HR, heart rate; MOI, mechanism of injury; SBP, systolic blood pressure; Temp, temperature.

In developing the model the authors undertook the following key steps: classifying coagulopathy, perfusion and tissue injury status; establishing the clinical relevance of coagulopathy; and developing an evidence framework which allows users of the model to explore the evidence underlying the causal structure (**Figure 13**). The model was trained using AgenaRisk software (Agena, London, UK).[323]

Classification of a coagulopathy is challenging as discussed in **1.4.4.2**. The authors performed a three-step process to classify patients with coagulopathy. First, they identified patients with an INR >1.2. Next, they separately performed patient clustering through AgenaRisk using an Expectation Maximisation Algorithm. This algorithm used all the available data points in the model to predict which coagulopathy group a patient belonged to. Data points included ROTEM measures of clotting function, lactate

measurements and the patient's injuries. Where the coagulopathy labels of the INR and clustering procedures agreed patients were given this label. Where there was disagreement between the two processes, three experts agreed on a label.

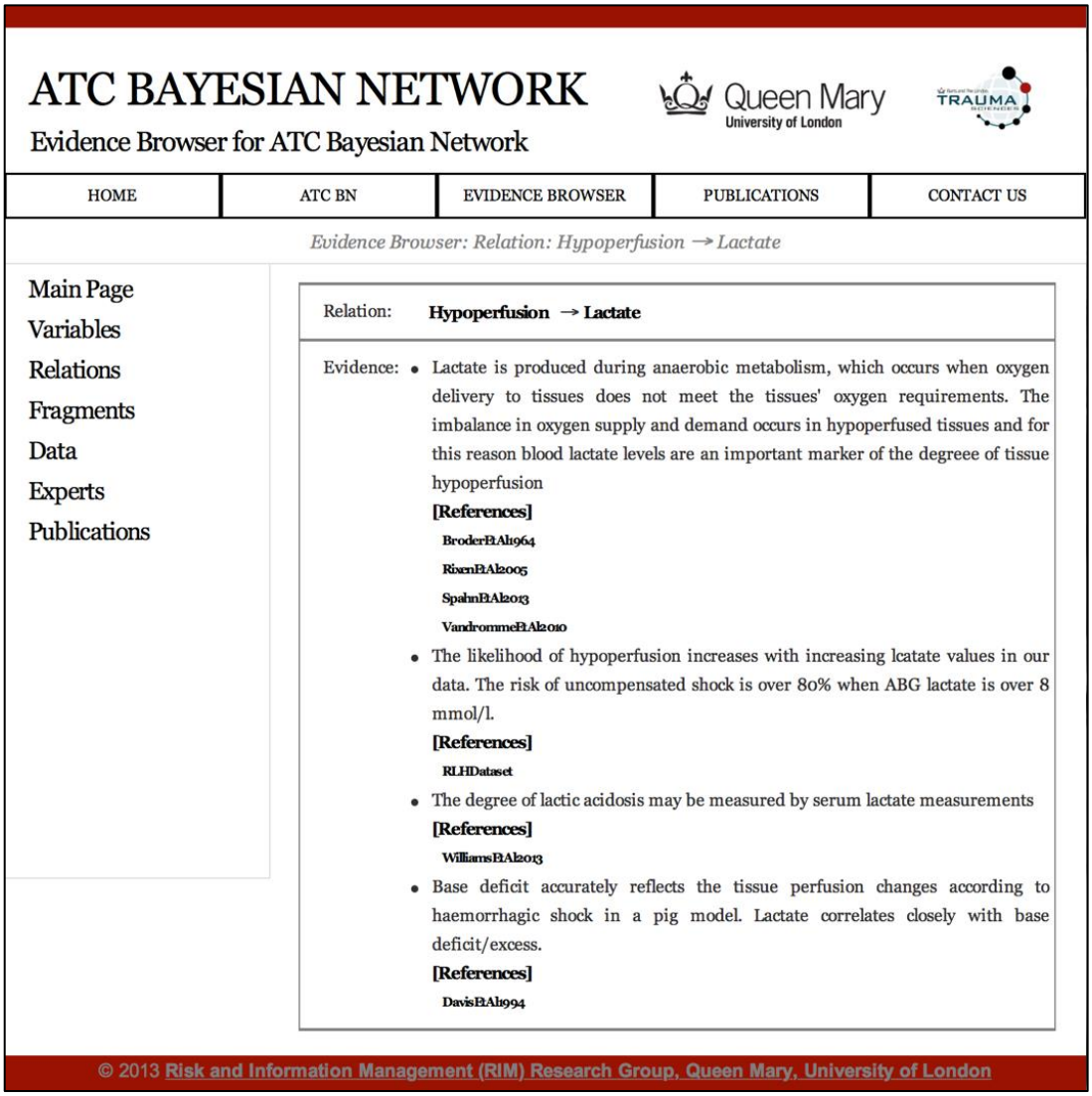

Figure 13: Example of the evidence browser developed by Yet accessible at www.traumamodels.com[340]

1.9.2 ED TIC BN predictor variable definitions

Domain knowledge (rather than a data driven process) was used to select predictors (evidence) for the causal mechanisms described in the model's DAG (**Figure 12**). The 14 variables were selected as they were known to be routinely generated during a standard ATLS primary survey (**Table 5**).

Table 5: Definitions and states of predictor variables in the ED TIC BN

Predictor Variable	Type of Node	Definition
Heart rate	Continuous	Heart rate, beats per minute
Systolic blood pressure	Continuous	Systolic blood pressure, mmHg
Temperature	Continuous	Body temperature, ^{o}C
Haemothorax	Boolean	Suspected based on examination or CXR findings
FAST result	Boolean	Intra-abdominal peritoneal fluid identified
Unstable pelvic fracture	Boolean	Suspected based on examination or PXR findings
Long bone fracture	Boolean	Suspected fracture of femur, tibia or humerus or traumatic amputation proximal to ankle or elbow
Glasgow coma score	Ordinal	GCS on admission or before intubation
Lactate	Continuous	Admission arterial or venous blood gas analysis
Base deficit	Continuous	Admission arterial or venous blood gas analysis
pH	Continuous	Admission arterial or venous blood gas analysis
Mechanism of Injury	Boolean	Blunt or penetrating
Energy of injuring mechanism	Boolean	High or low energy[1]
Volume of fluid administered	Continuous	Volume of crystalloid or colloid fluid given, mL

FAST, Focused abdominal sonography in trauma; CXR, Chest X-ray; PXR, Pelvic X-ray; GCS, Glasgow coma score

[1]The energy associated with injuring mechanism was defined as

1) high energy: high-velocity GSW; fall >20 feet (6 m); pedestrian or cyclist versus vehicle >20 mph; road traffic collision with mechanical entrapment, ejection from vehicle or death in same passenger compartment; entrapment under a train or vehicle; crush injury; blast injury
2) low energy: stab; low-velocity GSW; and blunt injury excluding injuries above

1.9.3 **ED TIC BN model validation**

The ED TIC BN model was validated using an external dataset of 491 patients. The patients used for external validation were separated by time from the development patients (373 patients) and two different hospitals (118 patients). Performance of the model was accessed in three main domains; discrimination using AUROC, calibration using graphical description and accuracy using the Brier score. In external validation the BN had excellent prognostic performance (**Figure 14**), was accurate (Brier Score 0.06 (0.05-0.08), and well calibrated (**Figure 15**). The BN provided greater accuracy than any individual predicting factor and had comparable accuracy to laboratory tests of coagulation but within shorter timescales. In addition, the BN is able to maintain the accuracy of its predictions with up to one third of predictor information missing.[337]

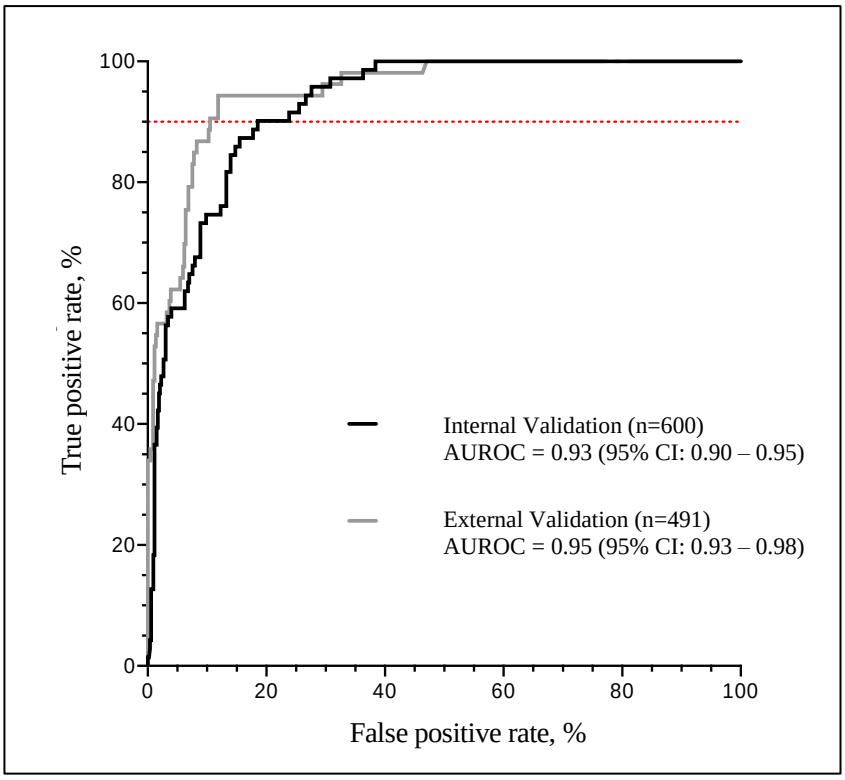

Figure 14: ROC curves for the validation of the prognostic performance of the Perkins-Yet ED TIC BN in patients recruited to the ACIT study. ROC curves show the relationships between true-positive and false-positive TIC predictions. The red dotted line marks the recommended operating point of the prediction at 90% sensitivity.[337]

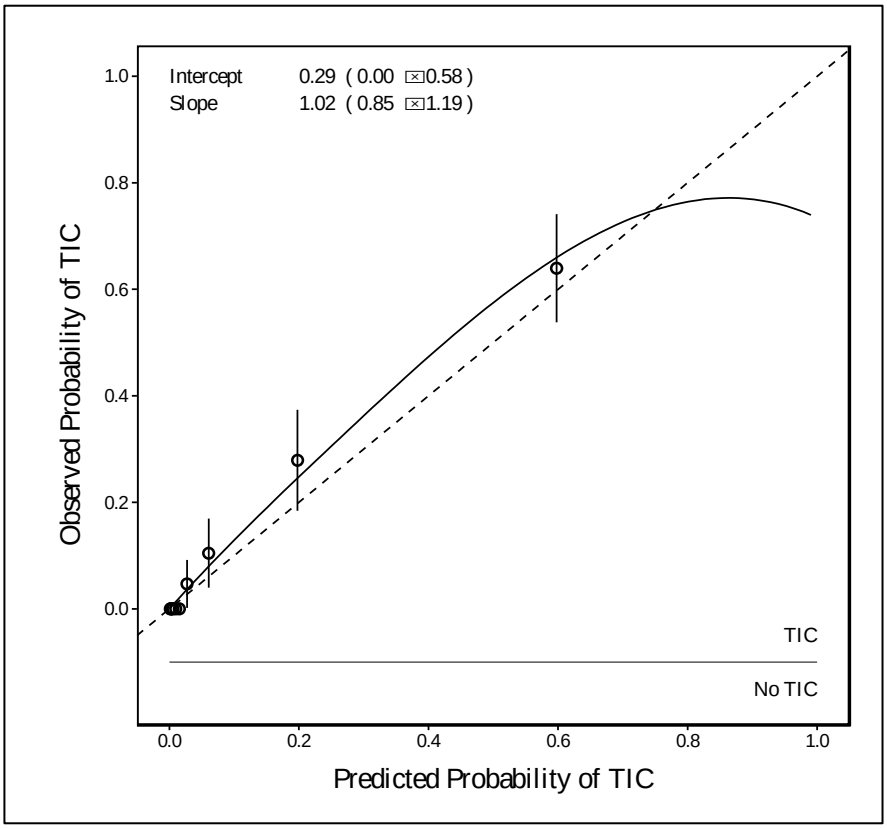

Figure 15: The ED TIC BN calibration plot. The plot shows the relationship between ideal (dashed line) and observed (solid line) predicted values in the development cohort. The rug plot along the x axis denotes observation frequency. The circles with 95% confidence intervals indicate observed frequencies by decile of predicted probability in 10% bounds of predicted risk. Error bars denote the 95% CI.

1.9.4 **ED TIC BN application**

The authors rendered an interactive interface of the model at www.traumamodels.com. With this interface users can enter patient information to compute the patient's risk of TIC. The graphical output of the model was not optimised at the time of hosting the webpage. Nevertheless, the model is usable and accessible by anyone with internet access.

1.9.5 **ED TIC BN impact analyses**

Mossadegh and Kyrimi describe their research to assess the impact of the ED TIC BN in their respective PhD theses.[327, 341] The research aimed to assess the performance of the ED TIC BN in real world conditions. The assessment was approached in three graduated steps.

First, Mossadegh demonstrated that all 14 of the input variables for the model were available with a median of 15 minutes (IQR 12-20 minutes) in a series of 67 prospectively observed trauma responses. Additionally, all data, except the blood gas measurements, were available within 6 minutes (IQR 4-11 minutes). Predictions generated at an earlier time, without the blood gas variables, were highly correlated ($\rho = 0.944$) to the final prediction with the full dataset.

The authors subsequently performed a series of experiments to assess the impact of the ED TIC BN on decision making. Initially they assessed hypothetical decision making and followed this up with a prospective real-world evaluation. In the first experiment, they conducted interviews with clinicians using 10 standardised clinical vignettes. The study was designed to assess the impact of decision making with and without the ED TIC BN. They enrolled 32 clinicians and demonstrated that use of the ED TIC BN appropriately reduced risk prediction when coagulopathy was not present. The study demonstrated that trust in the tool was associated with clinical experience. Junior staff were more likely to trust the tool than senior clinicians.

In the second part of their study they prospectively assessed the impact of the ED TIC BN on real world decision making. Nested within this study, the authors also performed a prospective external validation to measure the tool's performance. For this study researchers observed the initial management of 61 trauma patients in the ED. Ten minutes after the patient arrived, researchers asked the trauma team leader to give their prediction of coagulopathy. Concurrently, the ED TIC BN's prediction was calculated. The model's prediction was initially blinded from the trauma team leader, until after the patient had left their care. Blinding of model's prediction was performed to ensure the

model's prediction could not influence care, as the impact of the tool is incompletely understood.

Once the patient had left the ED, the trauma team leader's 10-minute risk prediction was compared to the model's computed prediction. Importantly, and similar to the clinical vignette study, the results showed that addition of the ED TIC BN result appropriately reduced coagulopathy prediction when no coagulopathy was present. In patients with coagulopathy the model did not alter the already appropriately high clinical predictions.[341] In addition in this relatively small sample, the external validation demonstrated high discriminatory performance (AUROC 0.98, 95% CI 0.95–1.0).

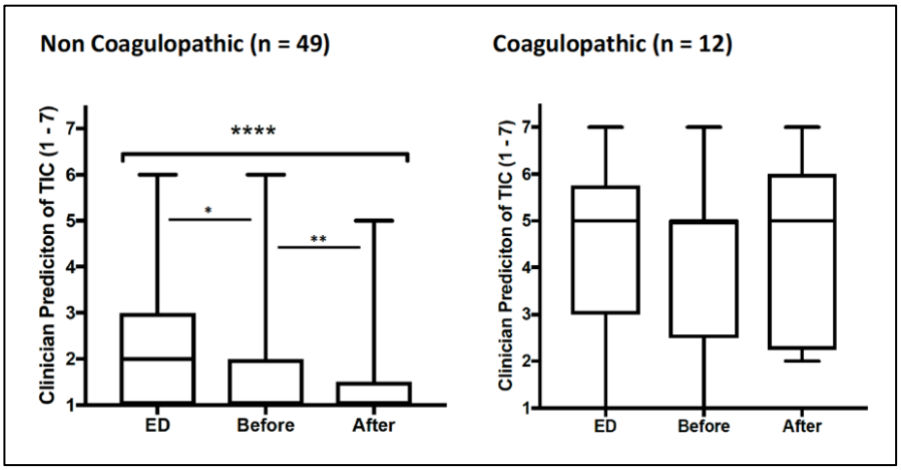

Figure 16: Boxplots demonstrating clinician's prediction of coagulopathy at three time points. ED; 10 minutes into the trauma call, Before; before the clinician was told the ED TIC BN prediction once the patient had left ED, After; after the clinician was told the ED TIC BN prediction once the patient had left ED. Y axis: clinical prediction quantified using a 7-point Likert scale. Left figure; patients without TIC. Right figure; patients with TIC. Figure from Mossadegh's PhD Thesis.[341]

1.9.6 Summary of the development, validation and preliminary impact analysis of the ED TIC BN

The ED TIC BN is a hand built causal probabilistic model trained and validated on highly structured data from a prospective cohort study. The strong performance metrics in retrospective internal-external validation, were reproduced in a small real-world prospective external validation. As the model reliably produces results long before

laboratory diagnostics, it has the potential to influence decision making in the ED. Encouragingly, initial impact studies suggest the model is able to support clinician's assessment of risk.

1.9.7 **What remains unknown**

Despite the model's promise, several aspects of its ability to support decisions remain unknown. For example, it is not known what proportion of trauma patients pose difficult decisions and may benefit from decision support. Nor, whether clinicians will be able to integrate the model's output into their decisions. Ultimately, for the BN model to have clinical impact clinicians will need to adjust their decision making and model augmented decisions need to improve health outcomes. To address these questions, sequential impact analysis is required. Early impact analysis will ascertain important information such as the size of any effect of the model on health outcomes. More definitive impact analysis requires an RCT of trauma patients treated with and without the assistance of the model. Until the RCT is performed, the full impact of the model on clinical care remains unknown.

It is well understood that in patients with major traumatic injuries, early intervention can be lifesaving (see **1.4.3**). It is not known, however, whether decision support used earlier after injury could have a more significant impact than in-hospital decision support. Nor, whether a novel pre-hospital BN model could retain the predictive performance of the ED model. Practical considerations about whether clinicians can readily capture a pre-hospital model's inputs and integrate the model's output into their decisions need addressing before impact studies can be considered. Finally, like the ED TIC BN, the impact of a PH TIC BN would need detailed assessment in an RCT randomising clinicians to the model or not. The definitive trial's endpoints should assess patient outcomes, clinician decision making and health system considerations such as resource use, and cost.

1.10 Aims and objectives of thesis

The aim of this thesis is to assess the impact of Bayesian Networks on decision making. The thesis focuses on decisions in pre-hospital trauma relating to major haemorrhage and trauma induced coagulopathy. The thesis is divided into the following aims:

- Describe how pre-hospital clinicians make decisions relating to bleeding and coagulopathy after injury and understand the conditions that may make decisions difficult. (CHAPTER 2, p84)
- Assess the potential impact of the ED TIC BN to improve recognition of major haemorrhage (CHAPTER 3, p120)
- Develop and assess the performance of a BN for pre-hospital TIC and major haemorrhage prediction (CHAPTER 4, p145)
- Assess the impact of the pre-hospital BN on the clinical assessment of TIC and major haemorrhage (CHAPTER 5, p169)
- Understand how uncertainty in clinical diagnosis affects modelling performance (CHAPTER 6, p206)
- Systematically review the knowledge that determines the casual relationships of the TIC BN (CHAPTER 7, p216)

1.11 Summary

Trauma is a leading cause of death in young people, and bleeding is the leading cause of preventable death. Haemostatic dysfunction is common after injury, and patients with coagulopathy have high rates of morbidity and mortality. Coagulopathic patients require more blood products, surgery, and longer in critical care.[342-344] There are therapeutic strategies for TIC, and the earlier these interventions are applied, the greater the benefit.[92, 345-347] Clinicians' judgement is needed to identify patients that may benefit from high-risk treatment strategies. The assessment of this risk is especially difficult in the pre-hospital environment, which lacks the accurate diagnostic adjuncts found in hospitals. There are no reliable PH diagnostic tests for TIC. Simple scoring systems to predict TIC have been developed but are either methodologically flawed[255] or not accurate enough to be clinically useful.[257, 261, 266] Current practice, therefore, relies on blind, empirical protocols and expert opinion.

Many decision support tools have been developed for the diagnosis of major haemorrhage and TIC, but very few have been validated, and none have undergone prospective impact analysis. Improvements in computing power and data collection have enabled the development of powerful machine learning techniques. Bayesian Networks are a machine learning implementation of Bayes Theorem. By using sound probabilistic laws, BNs can combine knowledge and data. BNs could be useful for pre-hospital clinical decision support as they reason under uncertainty, and their predictions are explainable.

A BN to predict TIC has been developed, validated and undergone initial impact assessment. The ED TIC BN is an accurate, well calibrated model that can adjust clinician's prediction of risk. The extent to which the ED model is generalisable to the pre-hospital environment is unknown. Further research is required to develop a PH TIC BN, assess a novel model's ability to alter decisions, and improve outcomes.

The next chapter explores how decisions about major bleeding and TIC are currently made pre-hospital without decision support.

CHAPTER 2

PRE-HOSPITAL DECISION MAKING IN TRAUMA: DESCRIBING THE CURRENT PROCESSES AND THE CHALLENGES

2.1 Scope of the chapter

This chapter explores how decisions are made and what challenges easy decision-making in pre-hospital trauma care. An interview study of expert pre-hospital clinicians is used to understand how decisions about bleeding and trauma induced coagulopathy are made. In depth exploration and understanding of the issues are elicited using a qualitative methodology. The chapter highlights the significance of clinical uncertainty and discusses both the opportunities and challenges for pre-hospital clinical decision support.

2.2 Introduction

Decisions in pre-hospital trauma care are characterised by rapidly changing circumstances, high levels of uncertainty and high stakes consequences. To make a decision, clinicians assess multiple sources of information. The available information differs in accuracy, reliability and usefulness. When information is uncertain, conflicting or missing, decisions are difficult and the ability to logically compare the utility of alternative courses of action is lost.

Trauma induced coagulopathy (TIC) and major haemorrhage are two interconnected phenomena that occur in severely injured trauma patients.[50] There is an increasing number of effective management strategies for both conditions. These strategies are often only effective when high-risk patients are identified early and rapidly treated.[142, 348] (see **1.4**). Clinical identification of high-risk trauma patients *in-hospital* is difficult and has been demonstrated to have poor accuracy.[99]

An observational study in 2015, demonstrated that trauma surgeons at 10 level 1 trauma centres in the USA were unable to accurately identify patients with major haemorrhage. In 966 patients that received at least one unit of blood transfusion, the surgeons had a 66% sensitivity and 64% specificity. The surgeons falsely predicted two incorrect major haemorrhage patients for every correct prediction. The surgeons were more likely to predict major bleeding in patients who were younger, had sustained penetrating trauma, were more severely injured, had higher heart rates, and had lower systolic blood pressures.[99] These findings are important to this thesis for several reasons; first, it demonstrates the degree of uncertainty present in trauma; second, it highlights several patient factors that clinicians have learnt to associate with high risk. Third, it highlights that there is no equivalent pre-hospital evidence.

Despite over five decades of experience with clotting abnormalities after injury,[349, 350] identifying patients with TIC remains an uncertain clinical practice. There are no published studies that attempt to establish the clinical accuracy of the diagnosis of TIC, or pre-hospital blood transfusion. Relative to *in-hospital* decisions, *pre-hospital* decisions are complicated by having fewer diagnostic aids, challenging environmental

conditions and less time from injury to assessment for the patient's physiological response to evolve.

The study presented in this chapter seeks to understand the factors that make decisions difficult among pre-hospital clinicians. The limitations of the classical approach to the diagnosis of haemorrhage is discussed in section **1.4.4.1**. What makes decisions around the diagnosis and treatment of major bleeding and TIC particularly challenging in the pre-hospital environment is not well understood. Increasing our understanding of the causes of difficulty, provides an opportunity to evaluate where an intervention to support clinical decision making, such as a decision-support model may add value and how it may integrate into the clinical workflow.

This study seeks to identify the challenges that experienced pre-hospital doctors face in making critical decisions and identify barriers to the integration of a decision-support tool. The study will provide an assessment of the pre-hospital *deployment terrain* and the processes and culture which may potentially affect decision-support integration.

2.2.1 **Qualitative Research**

The analytical method in this chapter originates from a distinct philosophical position in contrast to latter chapters within the thesis. Historically, the sciences are conducted in a positivist and post-positivist theoretical framework with quantitative methods. Increasingly, it is recognised that some questions are better suited to qualitative analytical approaches which have their foundations in a constructivist epistemology. Research questions that concern *"the meaning which people attach to actions, decisions, beliefs and values"*[351] are the type of questions that can be addressed with a constructivist or interpretivist approach. Data gathered from questions about human thoughts are commonly narrative and contain considerable complexity. Qualitative analytical approaches can provide an in-depth insight into this type of data.[352] Accordingly, a qualitative approach is used in this chapter to derive detailed and interpreted insights into how experts make decisions.

2.3 Aim

The two main aims of this chapter were to understand how pre-hospital decisions about bleeding, coagulopathy and transfusion are made by expert clinicians and additionally, what makes these decisions difficult. A secondary aim was to explore how probabilistic decision support may integrate into the pre-hospital clinical environment.

2.4 Methods

2.4.1 Reporting guidelines

Enhancing the QUAlity and Transparency Of health Research (EQUATOR) network (http://www.equator-network.org) is an international organisation to improve the quality and transparency of research. EQUATOR compile a database of reporting guidelines to help further this aim. For qualitative studies, EQUATOR's most recently recommended guideline is the SRQR.[353] SRQR consists of 21 standards to guide the reporting of qualitative research, which was followed to facilitate critical appraisal of the study.

2.4.2 Qualitative approach and research paradigm

This study poses the follow questions:

1) How do pre-hospital clinicians make decisions about bleeding, transfusion and TIC?
2) What makes these decisions difficult?
3) How might probabilistic decision support integrate into the pre-hospital clinical environment?

To answer these questions requires a qualitative analytical approach. There is no single accepted way of performing qualitative research. Multiple methodological frameworks exist within qualitative research. Well-known examples include Grounded Theory

which assumes the researcher has no pre-conceived ideas on the research area and might be employed to identify underlying theoretical explanations. Phenomenology is another well described method which can be used to help to discover abstracted essences of specific lived experiences. Some authorities suggest quality in qualitative research is achieved by consistently adhering to the philosophical starting point of a given methodology.[351] Equally, other authors encourage methods associated with a range of philosophical positions. These authors believe the quality of a study is enhanced by taking multiple philosophies.[354]

The design and analysis of this study, uses a Qualitative description approach.[355-357] The aim of qualitative description research is to provide a detailed description of the concepts under investigation in an easily understood language.[358] This approach does not constrain the analysis to one particular paradigm or guiding theory.[359] As the approach is free from any one theoretical framework, it therefore avoids inappropriate theoretical complexity. Broadly speaking, qualitative description approach focuses the analysis on the description given by the participants and the research "strives to stay close to the surface of the data".[357] The approach uses a combination of literal descriptions and is coupled with analysis which involves interpretation of the participants ascribed meaning.[360] In an applied health services research context, qualitative description provides a method of addressing the specific *a priori* research questions while also allowing for de novo data to arise and be incorporated.

In a qualitative descriptive approach, the role of the researcher in the interpreting the data is specifically acknowledged. It is accepted that there are many interpretations of reality. The output of this qualitative analysis is a subjective interpretation which is strengthened and supported by reference to verbatim quotations from participants.[360] The study design does not seek a replicable objective reality but rather an account that most of the researchers and participants involved can agree on.[358] The ultimate aim is to gain an "inside or emic knowledge" and learn how the participants "see their world".[360]

In this study, an empathic neutrality position has been taken. This empathic neutrality approach acknowledges that insights from the researcher will be used during data interpretation but maintains a non-judgemental stance.[351]

2.4.3 **Researcher characteristics**

In qualitative studies, the characteristics of the researcher, influence the research findings. *Reflexivity* is the term given to this circular relationship between cause and effect.[351] In this study, a stance of "empathic neutrality" was adopted. Empathic neutrality recognises that there is value-mediated by the researcher on the findings of the study. The empathic neutrality approach aims to make assumptions explicit. For the purpose of making assumptions clear, the researcher's characteristics, education and relationships with the participants are described.

I am a white British male. I attended Medical School in London and am now a General Surgery Registrar training in London. I have basic pre-hospital emergency experience in a military context and no subspecialty pre-hospital training in a civilian context. I conducted this study as part of a doctoral research degree at the Centre for Trauma Sciences, Blizard Institute, Queen Mary, University of London. I am employed by the UK Ministry of Defence. There are no other sources of funding for this study. All the participants in the study are known to me on a professional basis. I am influenced by the works of Kahneman (Heuristics and Biases, HB) and Klein (Naturalistic Decision Making, NDM) in the critical analysis of decision making.

In the course of this study, I was advised by Colonel Nigel Tai; Consultant Trauma Surgeon and my primary PhD supervisor, Professor Julia Williams; academic paramedic with interests in pre-hospital research and qualitative methodology and Dr Suzie Kellet; an anaesthetist with an interest in decision making. I was assisted in the transcription of the interviews and initial data analysis stages by Dr Rahul Bagga, an academic junior doctor.

2.4.4 **Study Setting**

The study was conducted at two Air Ambulance sites; London's Air Ambulance (LAA) and Air Ambulance Kent, Surrey and Sussex (AAKSS). Both organisations are in the South of England and provide a constant physician and paramedic response to critically injured patients. While there are many similarities between the two services there are

also some important differences. LAA provides care within an urban major trauma network whereas AAKSS serves a larger geographic area with a lower population density and greater distances to major trauma centres. Patients treated by LAA have a relatively shorter pre-hospital phase of care and are more commonly transported by road rather than air. LAA typically treats a younger population more frequently injured by a penetrating mechanism injury.

2.4.5 **Sampling strategy and participants**

The study used a purposive sampling strategy.[360] Purposive sampling is a non-probability sampling strategy in which the researcher selects participants that they believe will have the necessary experience to answer the study question. Purposeful sampling of a highly informed group is known to provide "rich, accurate and helpful information" about a specific question.[361]

Currently practising pre-hospital consultants of at least five years continuous pre-hospital experience were approached from the two participating organisations; LAA and AAKSS. The researcher invited participants that were known to have an academic interest in haemorrhage or TIC and are thus considered experts. Participants were not selected on the basis of their interest in decision making. Participants were invited to participate by e-mail. Participants were selected for interview based on their availability during the sampling period. No invitations were declined. The purpose of sampling participants from two separate institutions was to understand commonalities as well as differences between both the individuals and the institutions. Participants were given a number to maintain their anonymity and the sites will now be referred to as site A and site B.

2.4.6 **Ethical issues pertaining to human subjects**

Approval for the study was granted by the Health Research Authority as part of the EmPHATTIC study; IRAS 200742 and was reviewed and sponsored by Queen Mary, University of London. The EmPHATTIC study did not require REC approval. A

participant information leaflet and a consent form were given to all participants and written consent was obtained before the start of the study.

2.4.7 **Data collection methods, instruments and technologies**

The study used semi-structured interviews.[362] Interviews were conducted between December 2018 and January 2019. The interview guide and questions are found in Appendix A. The interview questions were informed by discussion between the researcher and the primary supervisor. All of the interviews were conducted face-to-face at the Air Ambulance headquarters by the primary researcher. The interviews were electronically recorded using a Snowball ICE microphone (Blue Microphones, Newark, USA) to the voice memo application (version 2.0 Apple Inc., Cupertino, USA). Interview audio recordings and transcripts were stored on a password-protected computer.

During and immediately after the interviews, notes were made on the immediate impressions of the participants' responses. These notes were referred back to in the analysis phase. Notes were also made during the transcription and analysis phases, which informed the formation of themes.

The recordings were transcribed verbatim. Transcription was achieved in a two-step process; first, the audio file was uploaded to Trint (Trint Ltd, London, UK) a commercially available voice-to-text software. In the second step, the automated text transcript was checked for errors by two researchers (MM and RB) using the original audio recording. The checked transcripts were imported to NVivo Version 12 for Mac (QSR International Pty Ltd, Doncaster, Australia), a data management software, to facilitate data analysis.

2.4.8 **Data processing and analysis**

Data processing and analysis was approached in six phases.[363] Phase 1 was data familiarisation. This occurred through reflection of notes and re-reading of the interview transcripts. Phase 2 was the inductive generation of initial codes within the interview

transcripts. Initial open data coding was performed independently by two coders (MM and RB). Open coding was followed iteratively by focused coding to refine ideas and develop themes. An iterative process of constant comparison for code development and refinement was employed.

In this context a *code* identifies a feature of interest within the data that is the *"most basic segment, or element, of the raw data or information that can be assessed in a meaningful way regarding the phenomenon"*.[363] *Coding* is the process of ascribing codes to the interview transcripts and *coders* are the analysts that perform this task.

Phase 3 followed competition of the first cycle of iterative coding and a search for themes across the data set was performed.[364, 365] Initial themes were reviewed in phase 4 following discussion with RB and a third researcher (SK). In reviewing the initial themes attention was given to the proposed theme's internal homogeneity and external heterogeneity and the adequacy of the themes to answer the study questions. During this phase two additional themes were generated, two original themes were merged, and one theme refined and re-named. A subsequent round of coding took place where refinement of codes was achieved using a common codebook between coders. Differences of opinion relating to the most appropriate code for an interview excerpt was settled by discussion between coders. Where a difference in opinion remained the third coder made the final decision.[366] In phase 5 the themes were defined and named. Finally phase 6 involved writing the chapter.

Both thematic analysis and a count of the frequency of the codes were undertaken. Thematic analysis was performed first. Using a combination of thematic and frequency analysis facilitated an analysis which quantified, analysed and reported emergent patterns within the data.[363, 367] An analysis of the frequency of codes was performed by a simple count of each code in every interview. The result was tabulated under the heading 'frequency'. The number of interviews which the codes were identified in was recorded and tabulated under 'interview'.

2.4.9 **Techniques to enhance trustworthiness**

Respondent validation is a process that seeks to verify the accuracy of the data in the study. It is a process to ensure the researcher's interpretation of the interviews correctly reflects the participant's view.[351] In this study, transcribed interviews were returned to the participant with the coded text annotated. The interviewee was asked to validate the accuracy of transcription and raise any unintentional discrepancies in the coded interpretation of their interview. Participants were sent the first complete iteration of the codes without the broader themes. Participants were not asked to validate the final version of the coding structure.

2.5 Results

A total of 10 interviews were undertaken; four participants from site A and six from site B. There were nine male participants, reflecting the male preponderance of pre-hospital clinicians at both sites. Participants 6 and 7 were Anaesthetists and the other participants worked in Emergency Medicine. The average pre-hospital experience of participants was longer at site A (median 16 years site A vs. 9 years site B). Individual participant characteristics have not been detailed due to the risk of participant identification.

Interviews lasted a median of 30 (IQR 28-35) minutes. Once eight interviews had been conducted further data gathering provided diminishing returns. The final two interviews (interviews 9 and 10) did not provide additional themes or sub-themes. In addressing the aims of this study open coding of the interview transcripts produced over 140 initial codes. These codes were refined iteratively between the three coders, and overarching themes were generated inductively. The themes identified are recognition-primed analysis, uncertainty, imperfect decision analysis and clinician autonomy (**Figure 17**).

A frequency of codes analysis is provided in **tables 7-10**. These tables demonstrate the relationship between the data codes and the main themes as well as the frequency of theme, sub-theme and code reporting across the data set.

Before addressing *how* participants made decisions about bleeding and TIC, participants were first asked whether the decisions were important. All 10 participants agreed that making a diagnosis of major bleeding and the decision to transfuse blood were critical parts of their initial patient assessment. Only two participants said that they consciously considered whether a patient had TIC.

> *"[the assessment of bleeding] is a key consideration and really what my primary survey is trying to establish, in the pre-hospital setting, is identifying obvious signs of, or potential for, bleeding."* Dr 1

Accordingly, participants provided detailed descriptions of the overt analytical processes used to establish a diagnosis of major bleeding. Descriptions about when to start a blood transfusion and whether a patient was at risk of TIC were less precise and demonstrated lower levels of agreement between participants.

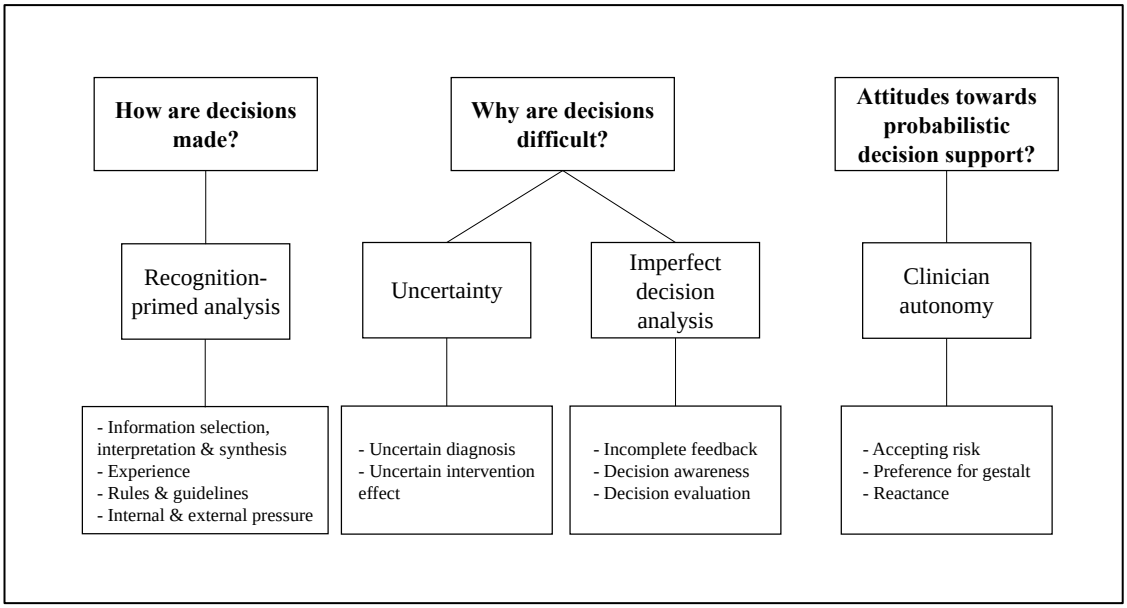

Figure 17: Connections between themes identified by thematic analysis. The top row of boxes denotes the three study questions, the middle row the overarching themes and the bottom row the subthemes.

Question 1: How do senior doctors make decisions about bleeding, coagulopathy and the need for blood transfusion in the pre-hospital environment?

When making decisions, it was clear that participants attend to information or cues from multiple sources. Many of the cues participants reported are well described elsewhere (see **1.4.4**). This section focuses on insights that are less well described.

The methods used to make a decision are complex and can be understood using concepts from Naturalistic Decision Making (NDM) (see **1.5.1.1**). Experienced clinicians used learnt patterns to select, interpret and synthesise key cues to make decisions. Decisions were also influenced by guidelines and psychological pressure. These influences were variable depending on the decision maker.

2.5.1 Theme 1: Recognition primed analysis

2.5.1.1 Information selection, interpretation & synthesis

Participants described seeking out familiar patterns starting from the initial job description.

> *"I mean part of the assessment probably starts before you even get to the scene ... knowing what job you're going to ... knowing your mechanism or what the alleged mechanism is that starts you thinking"* Dr 8

Once participants arrive at the patient's location, they use multiple cues to update their perception of the risk of major injury.

> *"A lot of the surrogates that we use to identify bleeding ... are not particularly sensitive and so you need to add as many layers to the picture as possible, it's almost like pieces of a puzzle that allows you to then stand back and look at the whole picture once you've got each piece"* Dr 6

Participants at both sites agreed some indicators were more predictive of life-threatening haemorrhage than others. One of the two sites teach their clinicians to identify a set of eight indicators they believe are strongly predictive of life-threatening haemorrhage. These factors are collectively referred to as the "hateful eight" (**Table 6**).

Table 6: **The hateful eight; factors suggestive of major haemorrhage**

Factor suggestive of major bleeding	Type of Factor	Source
Pale	Qualitative	Visual inspection
Clammy	Qualitative	Visual inspection / palpation
Apparent air-hunger (change in respiratory pattern)	Qualitative	Visual inspection
Venous collapse	Qualitative	Visual inspection
Hypotension (low volume or absent peripheral pulses)	Qualitative	Palpation
Low or falling end tidal CO_2	Quantitative / trend	Electronic monitoring
Tachycardia or relative bradycardia	Quantitative	Electronic monitoring
Altered mentation	Qualitative	Inspection

Cues relating to how the patient was injured influenced how participants interpreted the patient's injuries and the cause of abnormal physiological observations. It is this information interpretation that is termed here as Recognition-Primed *Analysis* (**Table 7**). This novel term describes an analytical step within RPD. First cues and information are gathered and then analysed in a primed-recognition phase before decisions are made. This analysis may have elements of System 2 cognition or may be automatic.

> *"doing the primary survey, I'd like to say that I'm always aware of the colour of their skin, and those sorts of things, but I'm probably not. Though sometimes it is obvious that someone looks very pale and you pick up those cues, but it's not always the case."* Dr 1

For example, recognition-primed analysis is used to avoid what participants called "haemorrhage mimics". Using the mechanism of injury, amongst other cues, primes the decision maker to search for specific injuries. This search is to avoid mis-diagnosing a patient as hypovolaemic due to bleeding. The search for alternative diagnoses is an analytical cognitive step.

> *"If you're to avoid an exsanguination mimic, you need to establish that they've had a mechanism of injury consistent with some injuries and you've found those injuries. Because you could have all of that physiology [hateful eight] … and not have any injuries… that's where it goes wrong; people just look at the physiology."* Dr 10

The use of point of care testing differed between the two sites. Participants at site A used point of care lactate measurement and ultrasound imaging to influence their decision making. Participants at site B did not use these tools and expressed uncertainty about their value.

> *"I really like the lactate, because I think it just gives me that extra dimension for those patients where I'm either on the fence or the patient doesn't, to my external assessment, declare themselves one way or the other."* Dr 3

> *"We don't currently do … blood gas testing but … I've always tried to think … if I could take the gas now what [would] the values be? … I hope that I transfuse patients who have a big base deficit or high lactate. … until the machines get a bit better I'm not sure they're going to add a lot to our clinical armamentarium."* Dr 5

These different views demonstrate differing methods to cope with uncertainty. Dr 3's view expresses the desire to gather more information. This strategy is well recognised within NDM (**Figure 5**). Whereas Dr 5's hypothesising relies on her experience and ability to match patterns. The quality of Dr 5's judgement is also dependent on the regularity of her environment and ability to learn within it.

The diagnosis and management of TIC was less clearly described by most participants. Some participants said they consciously assessed the patient's risk for TIC. These participants anchor their diagnosis of TIC on criteria such as the degree of tissue injury and duration of shock. Iatrogenic causes of coagulopathy such as dilution of clotting factors with crystalloids were also discussed. For the majority of participants that reported a non-formalised method of assessing TIC, the same broad contributing factors were reported.

> *"There are groups though where you think, you know, that it [TIC] is going to be a problem. … [for example] those patients with multiple tissue injuries… on scene it probably doesn't feature in what you're actually cerebrating, but you know that that patient is going to get into trouble, and getting into trouble, includes trauma induced coagulopathy."* Dr 8

Some participants felt the question of how to manage TIC pre-hospital was misdirected. These clinicians believed that TIC was a consequence of other processes such as

hypoperfusion and that the aim of management should be to minimise the factors that drive TIC. Many thought TIC was best managed in hospital with ready access to blood and blood products, whole blood coagulation diagnostics and surgical haemorrhage control. Thus, rapid transport into hospital was deemed beneficial.

2.5.1.2 Experience

The influence of experience was discussed throughout the interviews. Participants described a rapid intuitive decision-making process that was linked to having a body of experience. Experience changed the way clinical information was interpreted and synthesised. Commonly, participants described changes in opinion about when to start a blood transfusion.

> *"…one of the things that has evolved is that previously I would have been more swayed by physiology and perhaps less so by the findings on the primary survey and perhaps increasingly, I've moved slightly in the opposite direction"* Dr 2

Experience is necessary to appropriately interpret cues even when they have been highlighted as important such as the "hateful 8". These cues are likely to have considerable inter-individual variation in measurement.

> *"The breathing issue is something that unless you've seen a lot of people in this pre-hospital phase often by the time those patients get to hospital, they've declared themselves"* Dr 8

The diversity of prior clinical experience may also be responsible for differences in opinion between base specialities. An example here contrasts an Emergency Physician and an Anaesthetist both practicing pre-hospital medicine at the same Air Ambulance site:

> *"I'm comparing to my colleagues who are less cautious in giving blood products, I think that my threshold is relatively high to give blood products to these patients."* Dr 5

> *"I think I have less of less of an existential angst about giving blood than some of my colleagues."* Dr 6

2.5.1.3 Rules, guidelines and evidence

In pre-hospital medicine, standard operating procedures (SOPs) are regularly used to standardise medical provision. For pre-hospital blood transfusion, SOPs include criteria on when to start a transfusion. In discussing triggers for blood transfusion, most participants made reference to their SOP. This senior group of clinicians suggested that they were not rigidly adherent to SOPs.

Alongside the formal SOPs, several participants described loose self-generated 'rules' that influenced their decisions.

> *"If you've got a blood pressure that looks within a reasonable range (and essentially that is something in the 100 plus range) then it just lowers your clinical suspicion that this person has bled to a point that they're going to need blood products. It's not an absolute rule."* Dr 2

Other sources of information such as standard training courses were not referenced in the interviews. Academic research was frequently mentioned to support the participant's standpoint. Equally, the limitations of this evidence-base were acknowledged. Evidence was often described as insufficiently precise to assist in decision-making.

> *"We just still don't really know whether giving plasma is going to help them, whether packed cells are going to help… we know that these patients will be okay for a period of time, quite how long that is… I'm not sure anyone knows … there's good evidence for [permissive hypotension] for penetrating trauma, we know that… it's more difficult for blunt trauma."* Dr 4

2.5.1.4 Internal and external pressure

Four participants described how their decision making was affected by either internal pressure, or the anticipation of external post-hoc scrutiny. When asked how the presence of bleeding changes management one participant replied:

> *"I think it makes you more stressed. Sometimes it focuses you a bit more … but, whether you feel it or not, there's an increased level of stress."* Dr 9

Post-hoc scrutiny occurs during formal retrospective reviews of patient management. Anticipation of this clinical governance process appears to influence decision making for some participants.

> *"it seems to be almost a like a badge of honour for bringing [a patient] in and having avoided giving them the blood but actually their physiology is deranged, and they need the volume replacement"* Dr 6

Such comments suggest that the decision making amongst pre-hospital clinicians is not uniform. In some interviews there appeared to be a tension between what the participant believed to be best practice and what they believed their peers would construe as best practice post-hoc. Faced with this tension, some participants appeared to make to decisions to fit in with the wider group's expectations while others were content to do what they thought was right.

2.5.1.5 Section summary

This section addressed how experienced pre-hospital doctors make decisions about bleeding, blood transfusion and TIC. Decision making is a complex process that can be understood through the lens of NDM. The theme *recognition-primed analysis* is novel term to address a cognitive step before decisions are made. Patient situations provide cues which influence the search for familiarity and pattern recognition. Experienced practitioners believe that some cues are more important than others and have begun to teach these cues within clinical communities.

There were differences in opinion over the utility of point of care testing, the need to address TIC pre-hospital and blood transfusion thresholds. Guidelines and research evidence appeared to have little effect on influencing the decisions of senior decision makers. For some participants these decisions were made amongst feelings of stress.

Table 7: Codes relating to recognition-primed analysis theme

Sub-Theme	Code	Interviews, n	Frequency, n
Information selection, interpretation & synthesis	**Information selection (cues)**		
	Mechanism of injury	10	20
	Physical signs of injury (including hateful eight)	10	40
	Visual and non-quantitative signs of bleeding	9	43
	Quantitative signs and diagnostic devices suggestive of bleeding	3	6
	Using trends and anticipating trajectory	8	18
	Autonomic response to injury	5	9
	Factors associated with TIC	9	32
	Information interpretation		
	Develop a differential diagnosis	8	24
	External factors such as time and distance	6	16
	Individual patients respond differently	6	10
	Assessment of critical hypoperfusion	5	9
	Primacy of the primary survey	6	14
	Information synthesis		
	Combining information to make a decision	9	51
	Expectation prior to patient assessment	5	9
Experience	Influence of experience	9	34
	Interpretation of signs changes with experience	6	9
	Clinical culture	2	5
Rules, guidelines and evidence	Rule based decision making	7	16
	Example of following an SOP or protocol	5	12
Internal & external pressure	Internal pressures and stress	4	9
	Post hoc scrutiny	3	4

Question 2: What makes decisions about bleeding, coagulopathy and transfusion difficult for expert clinicians?

Difficult decisions were characterised by uncertainty and imperfect decision analysis. First, the influence of uncertainty in creating difficult decisions is discussed before moving on to insufficient analysis of decisions.

2.5.2 Theme 2: Uncertainty

All of the participants at one point in their interview, identified their own feelings of uncertainty as a barrier to making decisions. Their uncertainty could be split into two sub-themes. There was uncertainty surrounding both the patient's diagnosis and the benefit of a given intervention (**Table 8**). This combined uncertainty precluded effective prognostication, which in turn affected several key pre-hospital decisions.

2.5.2.1 Uncertain diagnosis

Participants reported situations in which they lacked confidence in the pre-hospital information. Most often this uncertainty was related to the inaccuracy of physiological observations to accurately portray the patient's true state.

> *"it would be lovely to remove some of the complicating factors or some of the unanswered questions and I guess, in my head, I'd love to know how well the [patient's] tissues are being perfused."* Dr 9

Patients injured with high energy blunt mechanisms often have multiple injuries and maybe lacking obvious cues such as external haemorrhage. Participants highlighted this group of patients as difficult to diagnose due to the uncertainty of the cause of their abnormal physiological observations.

> *"this guy had classical signs of bleeding according to the books and ATLS: tachycardia, low blood pressure, and pallor... But ultimately, he wasn't bleeding."* Dr 7

The diagnosis of TIC presented more uncertainty amongst participants than the diagnosis of major haemorrhage. The absence of pre-hospital TIC diagnostics was a commonly reported reason for an inability to make an accurate diagnosis. Participants also thought that clinical cues were less reliable for the diagnosis of TIC than major haemorrhage.

> *"I don't know how I would assess it [TIC] to be honest because there aren't the same clinical markers [as major bleeding]. You can hunch it, but if you were to say: is this person coagulopathic? Much more difficult to say with any degree of certainty, to objectively assess or make a judgement on that."* Dr 4

2.5.2.2 Uncertain intervention effect

Between participants there was variability on the intended outcome of pre-hospital blood transfusion. While most agreed on the short-term benefits, the longer-term benefits were more uncertain.

> *"I'm not sure anybody knows this as fact, but there's this worry that by having that period of having under resuscitated [a patient], you set in chain a load of things that are going to cause them difficulty in the future; be that coagulopathy or multi-organ dysfunction."* Dr 2

Participants agreed that for patients close to or in hypovolaemic cardiac arrest, a blood product transfusion should be started immediately. However, outside of this group, participants did not clearly articulate or agree on the point at which blood products should be given.

> *"[the benefit of transfusion] is balanced against not wanting to overshoot, push that pressure up, lose that clot and get active ongoing haemorrhage."* Dr 6

Patients that appeared to have a moderate, rather than immediately life threatening, degree of bleeding presented participants with difficult decisions. For these patients, participants were often unable to identify which patients would benefit from transfusion. As a result, there was a variable reliance on either gestalt or reversion to SOPs to make a decision. In patients where the benefit of a blood transfusion was uncertain,

participants agreed that the risks of under-transfusion were greater than the harms of an unnecessary transfusion.

There was more uncertainty amongst the participants about treating TIC. Some participants felt they had limited or no options to treat TIC pre-hospital.

> *"Outside the hospital, how does it [TIC] change management? Again, it doesn't really. I don't have the bits and pieces to a) evaluate properly if the patient has [TIC] and b) to more bespokely manage it."* Dr 6

Similar to the uncertain future benefit of transfusing hypotensive patients, was a concern that TIC may not be an active clinical problem in the pre-hospital environment.

> *"I think it's very hard to be motivated or clinically steered when something you're asking about we don't really see. … It may be argued that we could turn off [TIC] and that we should be turning it off earlier. If that's the case, then I think we would probably need some academic evidence that that coagulation [dysfunction] is present [pre-hospital] and we need to measure it to turn it off a bit later."* Dr 10

2.5.2.3 Section summary

In this section, uncertainty was identified as a central factor causing difficult decision making. Uncertainty originated from an inability to make a confident diagnosis, and not knowing whether an intervention would be beneficial for a particular patient. Uncertainty reduced effective situational awareness, prognostication and management planning. The section highlighted particularly challenging groups of patients and the uncertainty connected with using examination findings to diagnose major haemorrhage and TIC. Without specific diagnostics and therapies many participants felt it was not useful to consider TIC pre-hospital.

Table 8: Codes relating to uncertainty theme

Sub-Theme	Code	Interviews, n	Frequency, n
Uncertain diagnosis	Few pre-hospital diagnostic aids for major bleeding or TIC	10	31
	Haemorrhage mimics can confuse diagnosis (e.g. head injury)	8	22
	Difficult to prognosticate	8	19
	Need to compile information and make a global assessment	8	9
	Differentiating minor bleeding from major bleeding	7	13
	Haemodynamic assessment maybe misleading	6	17
	Fallibility of haemodynamic parameters	6	16
	Clinical examination is not 100% accurate	6	7
	Individual patient variability	5	15
	Clinical unknowns for your patient (e.g. extent of their injury)	4	12
	No clinical signs of TIC pre-hospital	3	4
Uncertain intervention effect	**Immediate effect of a given intervention**		
	Benefit of transfusion (includes clotting, oxygen delivery, perfusion, pre-load)	9	52
	Risks of transfusion (includes dilution, increased BP, immunological effects, BBV, VTE, metabolic)	5	17
	Weighing up benefit and harm	9	16
	TXA and haemorrhage control mitigates TIC	7	10
	Recognition of threshold variation between clinicians when to start transfusion	5	7
	Future effect of a given intervention		
	Transfusion mitigates future pathological states (e.g. TIC / MODS)	6	8
	Code red ensures skilled personal are there	6	11
	Blood transfusion improves patient's physiological state in moderate bleeding patients	3	3
	No perceived effect of a given intervention		
	Limited options to treat TIC pre-hospital	8	16
	TIC does not change initial management	7	12
	No evidence for treatment of TIC pre-hospital	4	8
	Rapid transport to hospital is key	4	6

BBV; Blood borne virus, BP; Blood Pressure, MODS; Multiple organ dysfunction syndrome, TIC; Trauma induced coagulopathy, VTE; Venous thromboembolism

2.5.3 **Theme 3: Imperfect Decision Analysis**

This section addresses the interviewees' awareness of decision-science. The section highlights examples of incomplete decision awareness, cognitive biases and the educational challenges present in a pre-hospital environment. It is proposed that decision making inattention contributes to difficult decisions by constraining effective learning.

2.5.3.1 Decision awareness

This experienced group of doctors were honest and reflective in their interviews. Perceived short comings in their own clinical abilities were discussed openly. In contrast, participants rarely discussed how they arrived at decisions. There were only three mentions of decision awareness throughout the interviews.

> *"… a lot of the assessment actually happens in the first few seconds and then there is the more formal examination and the rest of it follows… I think this is a bit about blink theory … experienced clinicians will make a judgment very very quickly … then you try to reinforce that with your full formal assessment of the patient, probably try to work out whether your gut reaction was the right one, because I guess, sometimes it isn't."* Dr 8

Another participant reflected on decisions that were made automatically and influenced by experience:

> *"… I think it's gut feel and because of that it's not always particularly sensitive and sometimes we get it wrong."* *Dr 6*

Both of the quotes reflect that decisions are influenced before all of the available information in processed. After which evidence is sought to re-enforce the initial "gut-reaction" which puts clinicians at risk of anchoring bias. The quotes also both suggest that the initial reaction is most often correct.

The final mention of decision awareness advocates for increasing conscious analysis to improve the accuracy of decisions.

> *"I don't think it's okay to say it all goes in, and it doesn't matter why we think they're shocked, but they're shocked. No, you have to consciously bring it to the front of your mind, and you will get it right more of the time.* Dr 9

This final quote demonstrates an understanding that decision-making could be developed with more attention. This view could be developed to enable clinicians to continue to consider the best option rather than stopping at the first option that could work.

Associated with the more common lack of attention to decision awareness, participants had difficulties articulating the factors which affected some of their decision-making. These difficulties were particularly overt when the decisions required a more nuanced balance of perceived risks and harms. For example, when participants were asked whether or not they would transfuse a patient when the benefit was uncertain, they unanimously agreed that the risks of under-transfusion were greater than the harm of exposure to an unnecessary transfusion.

> *"In the case of uncertainty… I'm going to give it [blood] to them... and I think that's because of our patient group. Most the time, the chance of them bleeding is higher than the chance of just having impact brain apnoea or an alternative, and I think also that we're quite good at excluding other causes."* Dr 5

Several decision awareness related problems are apparent from this quote. First, from an RPD point of view it is clear their experience influences their decision. If this is acknowledged they may be able to avoid the bias of generalising to the patient group rather than the individual, which is prone to error. Second, the belief that their decision making is quite good may not be accurate, and likely represents a confidence bias.

Throughout the interviews there are examples of deficits in decisions awareness. There were examples of decision making susceptible to biases of anchoring, availability, framing or substitution. Sometimes the risk of bias was explicitly stated but more often the bias did not seem apparent to the participant.

> *"Because we carry blood and because you've got a solution. You can make the patient fit your solution. But that patient may not be bleeding."* Dr 8

> *"I mean part of the assessment, really in your head, probably starts before you even get to scene because you, kind of, know what job you're going to or briefly what job you're going to I think."* Dr 8

Several participants demonstrated deferred decision making, possibly as a mechanism to cope with uncertainty. Typically, responses with deferred decision making advocated for therapy to start in hospital rather than pre-hospital. Conflation of two concepts was also noted.

> *"When you're looking at someone, and you're evaluating them for bleeding you're feeling the pulse, you're looking at the blood pressure, you're looking how sweaty they are, how pale they are, there's almost a checklist in your head that's ticking off little boxes as you go, whereas for coagulopathy, we've obviously got no good kit that gives us a clear answer."* Dr 6

This example appears to conflate increased uncertainty in the diagnosis of TIC compared to bleeding by the observation of clinical signs with the lack of point of care diagnostics. Initially the explanation appears to offer a rational basis for the decision. However, it more likely represents the increased comfort the clinician has for decisions about transfusion for bleeding than for TIC.

2.5.3.2 Decision evaluation and incomplete follow up

Effective decisions evaluation requires a shift in approach from evaluating a decision based on the clinical outcome, to evaluating that decision based on the information available at the time the decision was made. This includes recognising the various biases which may have influenced the decision at the time, regardless of the clinical outcome. In general terms, clinical medicine spends little time on effective decision analysis. This contributes to decisions that are difficult.

2.5.3.3 Section Summary

Contributing to the difficulty in decision making was imperfect decision analysis. The deficit in awareness of decision theory manifest as a difficulty in articulating decision reasoning. In turn this is making decisions more difficult for the participants. Numerous examples of decision making biases were present throughout the interviews.

Table 9: Codes relating to imperfect decision analysis theme

Sub-Theme	Code	Interviews, n	Frequency, n
Decision Awareness	Automatic decision making	7	20
	Evidence of framing bias	5	6
	Evidence of substitution bias	2	3
Incomplete feedback	The decision is judged on clinical outcome	2	2
Decision evaluation	Suggestion of flawed decision evaluation	3	5

Question 3: How might probabilistic decision support integrate into the pre-hospital clinical environment?

Participants were not directly asked how they believed decision support would integrate into the clinical environment. Instead, a series of questions regarding their view of probabilistic treatment thresholds was used to elicit their beliefs. The theme encapsulating the individual codes was called clinician autonomy and cognition.

2.5.4 Theme 4: Clinician autonomy and cognition

This section describes how threats to clinician autonomy, limitations of human cognition and beliefs about decision support create barriers to the introduction of probabilistic decision support. Sub-themes which describe the potential barriers were clinician's acceptance of risk, preference for Gestalt and a psychological phenomenon known as reactance.

2.5.4.1 Acceptance of risk

Participants expressed difficulties with trusting outputs from decision support instead believing their own intuitive decision processes were more reliable.

> *"if I have a test … that is 100 percent accurate … and I did that test and it comes out on the side of do something, I would be remiss not to give it due consideration. But because I'm a clinician there may be a one in 50 incidence where I think 'do you know what, I'm not going to do that.' But probably 49 times out 50, if the test supports my clinical picture then I will go along with that."* Dr 3

Associated with this, participants demonstrated significant cognitive exertion in their evaluation of probability and risk. Most participants asked for more clarity on the questions and several openly expressed the discomfort of uncertainty. Several participants were unable to articulate their decision-making processes in probabilistic terms. This difficulty may stem from a lack of insight into how heuristic based decisions are made. Not all participants displayed this discomfort and some participants were able to integrate hypothetical probabilistic decision support information into their decisions.

> *"when you think 'this patient is potentially not that bad' and actually your [probabilistic] model, with 80% or above certainty, says 'we've ruled this out, they're definitely not coagulopathic'. That would really change our decision making."* Dr 4

2.5.4.2 Preference for gestalt

During the interviews the majority participants expressed a preference for using their own heuristic approach to making a diagnosis. The concept of Gestalt originates from psychology and suggests that the nature of a unified whole is not understood by analysing its parts.[368] In healthcare the term gestalt refers to a heuristic approach by which a diagnosis and treatment plans are formed often within seconds.[99, 369] Commonly participants expressed views consistent with the belief that making a heuristic based decision was less risky. No participant discussed data, or the need for evidence, that measured the quality of their heuristic based decision making

2.5.4.3 Reactance

Within the context of the clinical uncertainty discussed in section **2.5.2**, participants were conflicted about how useful additional data would be. Several participants expressed the view that additional test results would make decisions more complex and

thus more difficult. Several factors may explain this view. First, adding an unfamiliar piece of information into an already complex decision will initially require additional cognitive processing. Second, some reactance may be occurring. Reactance is "an unpleasant motivational arousal that emerges when people experience a threat to or loss of their free behaviours."[370] Perhaps clinicians' perceive an erosion into their professional autonomy when requested to consider using a tool to augment their decision making.

2.5.4.4 Section summary

This section discussed three concepts that need to be considered for the introduction of probabilistic decision support into the clinical environment. Concerns regarding risk and trust, a preference for Gestalt and displays of reactance all represent potential barriers to the integration of decision support.

Table 10: Codes relating to clinician autonomy and cognition theme

Sub-Theme	Code	Interviews, n	Frequency, n
Acceptance of risk	Discussion of risk	8	13
	Preference to over triage rather than false negative	4	6
Preference for gestalt	Suggestion of over anticipation or pre-decision	9	14
	Evidence of preference for gestalt	8	24
	Gestalt versus tools (diagnostics or electronic decision support)	5	8
	Stated importance of first impressions	3	4
Reactance	Suggestion of adverse psychological reaction to clinical decision support	4	5

2.5.5 Summary of results

This study posed three research questions and generated four themes. The themes drew on NDM theory to explore how decisions are made, highlighted uncertainty and

imperfect decision analysis to describe what makes decisions difficult and finally reflected on a potential conflict between clinician autonomy and decision support.

Subthemes were used to group codes and to add detail to the analysis. To provide a broad overview, the subthemes have been summarised as either representing a challenge or an opportunity for the introduction of a CDST (**Table 11**). Allocation to either the challenge or an opportunity category was made by assessing the balance of data from the interviews. Where the data equally represented a challenge and an opportunity, both are indicated. Participants reported that identifying patients at risk of TIC was a challenge pre-hospital. The potential ability to quantify uncertainty around the diagnosis of TIC is therefore viewed as a strong opportunity.

2.6 Discussion

This study provides key insights into the context, opportunities and challenges for deployment of a decision support tool in pre-hospital trauma care. Detailed interviews with expert clinicians were explored using thematic analysis and four themes were identified. How pre-hospital clinicians make decisions is complex and involves the interpretation and synthesis of specific cues and experience. The recognition-primed analysis theme helps to understand this process. What makes decisions difficult is uncertainty and imperfect decision analysis. Not all management decisions are difficult for the bleeding patient. However, the management of TIC is shrouded in uncertainty. Finally, threats to clinician's autonomy, difficulty thinking probabilistically and participants' preference for gestalt are all identified as potential barriers to the introduction of probabilistic decision support. This section discusses the implications of the results of each question.

Table 11: Summary of challenges and opportunities for decision support to augment decision making in pre-hospital decisions about major bleeding and trauma induced coagulopathy

Themes and sub themes	Challenge	Opportunity	Explanation
Recognition-primed analysis			
Information selection, interpretation & synthesis		✓	A CDST will focus attention on key cues and effectively combine information. Expert cues can be integrated into the CDST
Experience	x	✓	Experienced clinicians maybe reluctant to incorporate advice from a CDST but if effectively integrated better decisions may be possible
Rules and guidelines		✓	CDST can be integrated into existing rules
Internal & external pressure		✓	Objective assessment provided by CDST can offload clinicians' emotional burden
Uncertainty			
Uncertain diagnosis		✓	Quantification of uncertainty is a key benefit of CDST. Especially relevant for TIC
Uncertain intervention effect	No effect		The current CDST has no influence on the effect of the intervention. Improved situational awareness may improve intervention selection
Imperfect decision analysis			
Decision Awareness	x	✓	Deficits in decision awareness and probability handling make the integration of a CDST challenging. The tool is likely to increase attention to decision awareness and this is likely to be beneficial
Incomplete feedback	No effect		CDST will not increase feedback
Decision evaluation		✓	A CDST prediction will allow decisions to be evaluated retrospectively with an objective anchor
Autonomy			
Acceptance of risk	x	✓	Clinicians may perceive their risk ownership changes either way with a CDST
Preference for gestalt	x		Conflicts with purpose of CDST
Reactance	x		Drives clinicians away from a CDST

2.6.1 How do senior doctors make decisions about bleeding, coagulopathy and the need for blood transfusion in the pre-hospital environment?

Decision making requires a complex process of information selection, interpretation and synthesis. In making decisions about bleeding and TIC, pre-hospital clinicians encounter significant complexity. They are forced to process large volumes of information. Much of that information is noisy and lacks specificity. To cut through this information fog, experts have identified cues used to make their judgements. The "hateful eight" is a set of cues that demystifies expert intuition and helps to transfer expert's tacit knowledge to other clinicians.[200] Although a useful process, some uncertainty remains. To the non-expert it is unclear how these pieces of information (or variables) should be prioritised during synthesis. Uncertainty arises from multiple interacting variables with intricate and non-linear relationships and a limit on the amount of information that can be kept in the working memory.[371]

Decisions made under uncertainty are often influenced by subconscious biases. Such decisions are prone to high degrees of variation between, and even within, individuals.[372] In this respect a prognostic model maybe able to improve decision making. Prognostic models combine multiple predictors to assess risk (see **1.5.3**). Models can be trained on previous patient data and if an appropriate model is used complex non-linear relationships can be learnt from the data. Using a model has the potential to rationalise the combination of variables and remove unhelpful human bias.

Identifying the experts' key cues is also important from a modelling perspective. If modelling techniques are used that are open to interrogation it is possible to compare the model's inputs with the expert's. This process of comparison is potentially useful to both parties. The model's validity can be assessed by experts and the expert's decision making can be examined by data processed through a model. Additionally, the use of a decision support tool can lead to better training by clearly defining what cues matter. If decision support can improve the quality of the priming and clinicians make the right decision more often, the CDST will increase the chance of learning the 'right' lesson from experience.

Examples from outside of healthcare demonstrate how decision support can be effectively used to focus attention to the information that matters. There are examples of model's useful application in high-risk industries such as military pilots[373] and nuclear power plant operators.[374] In an RPD framework the purpose of the decision support tool is to improve situational awareness and lead to better decisions.[375] This situational awareness revolves around the diagnosis of a problem or recognising when there is high risk of a problem occurring in the near future.

2.6.2 What makes decisions about bleeding, coagulopathy and transfusion difficult for expert clinicians?

Clinical uncertainty and the fallibility of human decision making are central to understanding the current ability of expert clinicians to make pre-hospital decisions. Uncertainty related to both the diagnosis and the benefit of a given treatment. In considering how decision support may integrate into the clinical environment it is important to note that the participants did not universally regard TIC diagnosis and treatment decisions as essential to pre-hospital practice. In contrast participants believed that judgements on major bleeding were essential. The apparent lack of prioritisation of TIC is not unexpected given the multiple uncertainties involved. Using Klein's RPD model (**Figure 5**) it can be seen that the four aspects of recognition fail in the context of decisions on TIC. There is both a lack of useful cues and uncertainty in regard to the plausible goals. Consequently, the actions to take are also uncertain. In this situation Klein advocates that the decision maker seeks more information.

Introducing a probabilistic decision support tool maybe one method to provide more information and quantify the uncertainty. Of course, at this stage, clinicians will still be left uncertain of the most appropriate course of action. Nevertheless, an accurate decision support tool will increase situational awareness. This in turn supports the correct interpretation of the clinical situation, selecting appropriate goals and ultimately increases the likelihood of selecting an appropriate action.[210]

While not specifically explored in the interviews, it can be inferred that clinicians experience significant cognitive exertion under conditions of uncertainty. One potential

benefit of a decision support tool is to cognitively de-burden clinicians. An accurate CDST may decrease cognitive effort by enabling the rapid synthesis of multiple cues. However, some participants were concerned that adding another data point from a CDST into the clinical milieu will inadvertently *increase* cognitive exertion. Both perspectives remain valid and require a well conducted trial to assess the impact of a model on cognitive load.

The theme imperfect decision analysis draws attention to a lack of emphasis on learning from decisions. The difficulty analysing decisions objectively is not unique to pre-hospital doctors. Often in medicine, retrospective review of activity as occurs in a morbidity and mortality meeting (M&M) can be judgemental rather than exploratory in nature. Such meetings are abound with inherent biases and often focus on the outcome rather than the decision.[376, 377] Examples of these biases are the natural tendency to avoid disagreeing with those we like and agree with those with authority.[378] Attendees at a M&M meeting may lack the ability, or knowledge of the available tools to assist in the construction of a narrative of events.[379-381] Without this narrative the decision points of a specific clinical management approach cannot be identified.

This awareness gap represents an opportunity to improve decision making quality. Training in decision analysis, may help develop insight. This in turn will enrich the case review process by more accurately and methodically identifying sources of risk and error. The use of standard tools and processes to support clinicians in evaluating their decisions would improve decision quality by improving the quality of priming and learning the 'right' decision from the experience.

In the present study the knowledge deficit in decision analysis is revealed in the difficulties the participants experienced in describing how they made their decisions. This issue has important consequences for the integration of new tools or ways of working. Specifically, it generates difficulty for participants to compare their heuristic decision to a probabilistic decision.

The perception that intuitive decisions (gestalt) are more reliable than probabilistic supported decisions, illustrates the difficulties inherent in incomplete understanding of

decision making. Lack of an analysis framework makes it difficult to create one's own internal database of the reliability of our decisions. The result is often a tendency to overestimate the reliability and quality of our decisions.[369] Thus, the validity of our intuitive judgments cannot be assessed by our subjective confidence.

In a collaboration between Kahneman and Klein, they described the environmental conditions required for effective learning from previous decisions. They posited that for intuitive judgements to be reliable there must be stable relationships between objectively identifiable cues and subsequent events.[200] Both the objectivity of identifiable cues and the ability to know the true subsequent event sequence are threatened in pre-hospital medicine. As a result, years of experience and effective learning is required to develop skilled intuitions. A CDST provides a method to faithfully combine cues and event outcomes to consistently assist this learning process.

2.6.3 How might probabilistic decision support integrate into the pre-hospital clinical environment?

The introduction of algorithms to replace, filter or assist human judgment often elicits a hostile response. The barriers to integration of decision support identified in this study have also been highlighted elsewhere. Challenges include issues with understanding, trust and the threat to autonomy.[382] Where practitioners feel more isolated, by virtue of working in smaller teams they may value their autonomy more than those working in larger teams.

2.6.4 Limitations

Within qualitative research, approaches to demonstrate the quality of the research and methodological rigor continue to be debated. The concepts of validity and reliability, familiar in quantitative studies, are commonly replaced with principles of credibility, dependability, confirmability and transferability in qualitative research.[383] These principles were developed in the 1980s and remain useful today.[360] The detailed description of the methods given at section **2.4** aim to address these four principles and

demonstrate the steps taken to ensure methodological rigor. The degree to which the researcher has influenced the analysis is left to the reader to decide. This critical appraisal is facilitated, in part, by the inclusion of the characteristics of the researcher (see **2.4.3**).

The study included 10 participants. Justifying the sample size in qualitative research is not possible using power calculations.[351] An argument runs that further interviews may lead to new themes. In this way, it is difficult to conclude that data saturation has occurred.[351] However, in the current study it became apparent that fewer new concepts were generated with the later interviews, such that each new interview provided few new codes and no further themes. In this respect the study reached inductive thematic data saturation.[384] Sampling was purposive although all the participants worked in Air Ambulance organisations that had a pre-hospital blood transfusion capability. Other Air Ambulance organisations in the UK are currently involved in a randomised controlled trial to assess the benefit of pre-hospital transfusion.[165] Clinicians in these organisations may have expressed differing views to the participants involved in this study.

The study's first question addresses how decisions are made about bleeding and TIC pre-hospital. To answer this question requires introspection of cognition. This is a challenging process as experts accumulate large bodies of knowledge though experience. Their perception and cognitive skills can be difficult to verbalise when they are not performing the task in a real-world environment. To address this Klein described a formal interview technique known as critical decision method (CDM) to elicit detailed information from experts using a series of probes.[385] The semi-structured interview technique undertaken in this study did not use the CDM method and future studies in this area may benefit from adopting that approach.

2.7 Conclusion

A vital starting point to develop a CDST has been to understand the current decision-making process and assess whether a problem exists. Pre-hospital clinicians make decisions on bleeding, transfusion and TIC which are recognition-primed thus heavily influenced by experience and subject to variable degrees of clinical uncertainty. Improved understanding of the decision-making processes has provided a theoretical perspective of how decision support tools may reduce some of the risks of bias.

The information derived from this study has provided key insights into the opportunities, challenges and role of decision support for major haemorrhage and TIC. The next chapter will examine the ability of a TIC clinical decision support tool to reduce clinical uncertainty and improve decision making.

CHAPTER 3

THE VALUE OF CLINICAL DECISION SUPPORT

3.1 Scope of the chapter

The purpose of this chapter is to understand the potential clinical value of a Bayesian Network TIC prediction model. In the previous chapter, difficult decisions were related to uncertainty and it appeared that clinicians favoured a heuristic decision-making approach. This chapter quantifies the predictive accuracy of native clinical judgement in the assessment of major haemorrhage protocol activation using a retrospective study design. The performance of a Bayesian Network (BN) decision support model is applied to the same patients. This comparison between clinical judgement and the BN model provides an insight into the model's potential value to augment clinical decisions. A sensitivity analysis is performed at two clinical thresholds of decision making.

3.2 Introduction

3.2.1 Impactful decision support

Clinical decision support tools (CDST), with prediction models at their core, have the potential to improve patient care.[199] Yet few CDST have proven useful enough to be

implemented in clinical practice (see **1.5.3.1**).[218] It is a long path from model development to improving patient care, in which multiple criteria must be satisfied (**Figure 18**).

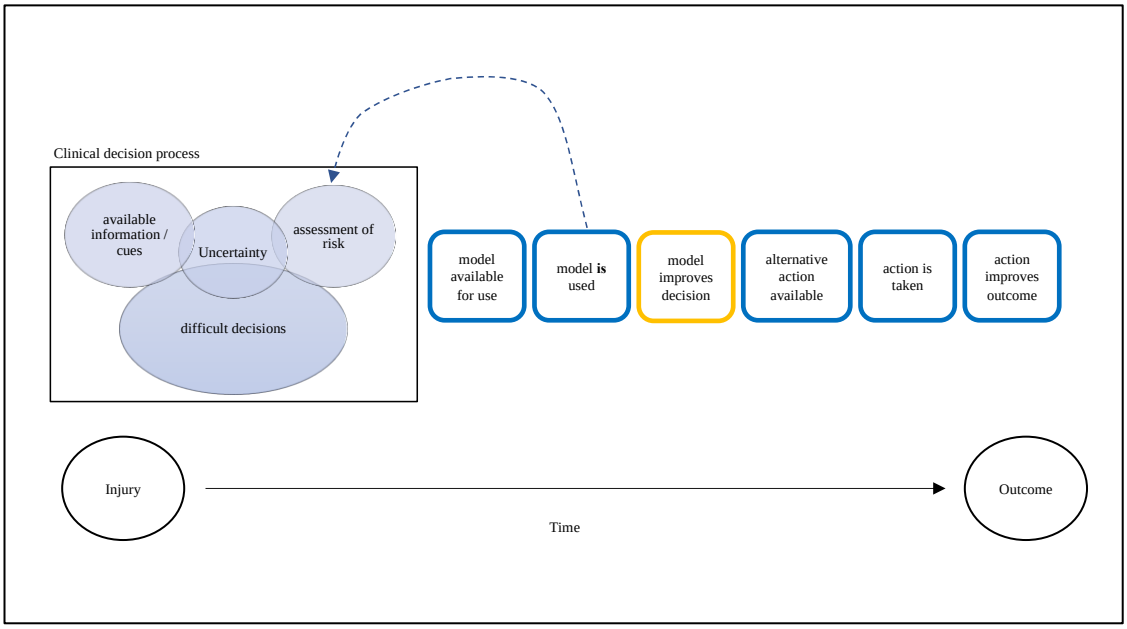

Figure 18: **Schematic of the critical steps required for a clinical decision support tool to improve patient outcome.** The figure demonstrates uncertainty identified in Chapter 2 following patient injury. This uncertainty contributes to difficult decisions. Each logical step (blue box) towards the outcome must be satisfied to ensure the tool has clinical impact. Dashed blue line denotes a change in the assessment of risk given the use of a prediction model. Enquiry into the step in the yellow box is the focus of this chapter.

Initially, decision support should be directed towards a clinical problem with uncertainty that stands to benefit from decision support. The degree of uncertainty associated with clinical decision making will, of course, vary considerably according to clinical context. Nonetheless those decisions where information is constrained by resource issues (lack of clinical experience, absence of sophisticated investigations or sufficient time to muster these adjuncts) are those where clinical decision support tools might add most value.

Subsequently, decision support must inform decision-making at a critical point in the treatment pathway. The CDST should readily enable the user to perform an evidence-

based analysis of individual patients. The CDST must improve the decision accuracy in the intended use population. Once the user receives the additional information, they must have the ability to decide between two or more actions. The user must act on the advice of the CDST, and that alternative action must improve an element of care. The benefit can be realised at a patient, treatment centre or health-system level.

Each link in the chain from model development to improved patient outcome (**Figure 18**) requires careful design and assessment. Some of these elements have well-established methodologies.[213,390] For example, external validation is a well-recognised step to assess the performance of a model in a population outside its development cohort.[205] Yet techniques to evaluate other steps such as the "model is used" and "action is taken" are not well established.

This chapter addresses the "model improves decision" step (yellow box **Figure 18**). To assess whether a model improves a decision requires an understanding of the baseline clinical performance.

3.2.2 Clinical use of bleeding clinical decision support models

Very few studies in major haemorrhage modelling address this performance gap between the tool and the clinician.[99, 233] Assessing the gap in prediction accuracy between clinicians and a prediction model is necessary to understand whether the model has the potential to improve clinical decision making.

There are over 40 published models to identify patients with major bleeding (see **1.6.1**). None of these models are compared to the performance of native clinical judgement during their development and validation. In the UK, bleeding decision support models are not in regular clinical practice. The NICE guidelines specifically warn against using these models.[110] In the USA, the ABC score is the most widely cited and endorsed for the prediction of massive haemorrhage.

3.2.3 **Assessment of the TIC BN to predict major haemorrhage**

The ED TIC BN model has been developed to risk stratify patients for TIC (see **1.9**). For TIC the model has been validated both internally and externally. In external validation, the model's excellent discrimination and calibration were maintained. TIC and major bleeding are interrelated in trauma. But it is not known to what extent the ED TIC BN could be used to predict patients at risk of bleeding. Furthermore, it is unknown if using the model to predict bleeding, rather than TIC, would have superior accuracy to native clinical decision making.

3.3 Aims and Hypotheses

3.3.1 **Aims**

3.3.1.1 *Aim 1:* Evaluate performance of clinicians to predict patient need for MHP activation.

3.3.1.2 *Aim 2:* Compare performance of the ED TIC BN to native clinical judgement to predict patient need for MHP activation.

3.3.1.3 *Aim 3:* Find the optimal operating thresholds for the ED TIC BN implemented to predict the need for MHP activation.

3.3.1.4 *Aim 4:* Compare the predictive performance of both native clinical judgement and the ED TIC BN to a well-known major haemorrhage prediction model.

3.3.2 **Hypotheses**

3.3.2.1 *Hypothesis 1:* Native clinical predictions regarding which patients warrant MHPA are inaccurate, and reduce decision making performance.

3.3.2.2 *Hypothesis 2:* The ED TIC BN has better predictive accuracy than native clinical judgement in identifying patients that required MHPA.

3.3.2.3 *Hypothesis 3:* The ED TIC BN and clinical judgement have better predictive accuracy than the ABC score in identifying patients that required MHPA.

3.4 Methods

3.4.1 Study Design

A retrospective study was performed comparing the real-world predictive accuracy of clinicians to a hypothetical decision made by the ED TIC BN. The original clinical decision for major haemorrhage protocol activation (MHPA) was identified in a sample of trauma patients enrolled in ACIT; a prospective observational cohort study.[339]

3.4.2 Patient sample

All patients were recruited to the ACIT study at the Royal London Hospital (RLH) between January 2008 and August 2013. Patients are eligible for inclusion in ACIT if they receive an in-hospital trauma team response, present directly from the scene of the injury, and are admitted within 2 hours of their injury. They are excluded if they are less than 16 years old, have burns more than 5% total body surface area, are anticoagulated, have a known bleeding diathesis or received over 2000mls of intravenous crystalloid pre-admission. Patients are retrospectively excluded if they decline consent. ACIT was approved by the National Research Ethics Committee of participating countries and written informed consent was obtained for all participants. Prediction modelling using patient data is a specific aim in the ACIT protocol and is thus granted ethical approval.

3.4.3 Data recording

For patients in ACIT, clinical management is conducted as usual with additional blood sampling on admission and throughout acute haemorrhage. The ACIT study prospectively collects multiple data points on each enrolled patient including patient demographics, mechanism of injury, admission vital signs, treatment administered and

outcome. Injuries were described using the Abbreviated Injury Scale (AIS)[386] and Injury Severity Score (ISS).[256] Patients are observed until hospital discharge or death.

3.4.4 **Study setting**

The Royal London Hospital (RLH) is a large urban academic Major Trauma Centre (MTC) in East London. The hospital functioned as an MTC with dedicated trauma resources throughout the study period.[22] During the study period activity at the hospital increased with an average of 1500-2000 trauma activations a year, of which approximately 25% had an ISS >15.[170] Trauma patients in the ACIT study arrive at RLH by one of three methods; i) self-presentation, ii) Ground Ambulance or iii) Air Ambulance. Clinicians on the Air Ambulance can activate the MHP pre-hospital.

3.4.5 **Clinical activation of the major haemorrhage protocol**

Senior physicians working at the RLH or LAA were responsible for MHPA. The clinical guideline for MHPA suggests activation with a systolic blood pressure less than 90mmHg, poor response to initial fluid resuscitation and suspicion or evidence of active haemorrhage.[170] These guidelines were not mandatory and clinicians could chose to activate the MHP if they believed the patient would benefit. MHPA could be requested either pre-hospital or in the emergency department (ED). Activation was recorded at any time until the patient left the ED. The time of MHPA was not recorded. Clinicians were blind to the model's predictions.

3.4.6 **The ED TIC BN model**

The development and validation of the ED TIC BN model are described in section **1.9**. The model uses 14 measurements identified in a standard ATLS primary survey to calculate a probability. In this study, the probabilities are reported as percentages. The 14 input variables are: mechanism of injury, energy of the injury, pre-hospital crystalloid volume, heart rate, systolic blood pressure, temperature and GCS on ED arrival, haemothorax suspected or identified on chest x-ray, abdominal bleeding suspected or

identified on FAST scan, long bone injury or an unstable pelvic fracture, lactate, base excess and pH (**Figure 12**).

3.4.6.1 From prediction model to decision rule

The output of the ED TIC BN is a probability. To facilitate a direct comparison of the model and the clinicians' decisions, the ED TIC BN prediction tool was converted into a decision rule. The tool's output was dichotomised by selecting an operating threshold. The operating threshold defines the value at which a decision is made. In this case the decision is whether or not to activate the MHP. A binary decision rule was consequently created converting TIC probability to MPHA.

No *a priori* assessment of the model's ability to predict MHPA was performed. The operating threshold of a 10% probability of TIC was chosen for MPH activation. This is the threshold used by Perkins and Yet to identify TIC with a 90% sensitivity (**Figure 14**). A *post hoc* sensitivity analysis was conducted in which the operating threshold of the ED TIC BN (i.e. the probability at which the model was dichotomised) was varied.

3.4.7 The Assessment of Blood Consumption (ABC) score

The ABC score is a popular massive haemorrhage prediction model (see **1.6.1.1**).[221] The model was derived from expert consensus and has four input variables. The presence of each is worth one point. The variables are penetrating mechanism, positive FAST, arrival SBP of 90 mm Hg or less, and arrival heart rate at least 120 bpm. A score of greater or equal to 2 is used as the operating threshold of this test. The ABC score has been included in this study to provide a comparison with clinical judgement and the ED TIC BN.

3.4.8 Appropriate major haemorrhage protocol activation definition

There is no recognised definition that characterises a decision to activate the MHP as appropriate or not. However, such a definition can be construed in terms of attributes related to patient outcome or indirect measures such as markers of therapeutic need (e.g.

metrics concerning burden of trauma or physiological disturbance). Need can also be inferred from a patient's consumption of therapy. This assumes that transfusion is accurately matched to a patient's need and patients are not over or under-transfused. Despite the risk inherent in such assumptions, there is good clinical concordance between the requirement for transfusion and likelihood of death in trauma patients. The risk of death increases in proportion to the degree of transfusion.

Stanworth *et al.* made the distinction between patients that receive at least 4 units PRBCs in 24 hours (major haemorrhage) and patients that receive 10 units PRBCs/24hours (massive haemorrhage). So defined, major haemorrhage patients have double the risk of death (25%) compared to patients who get less than four units, with massive haemorrhage patients exposed to a mortality risk of 33%.[387] In choosing markers that retrospectively denote appropriate activation of a major haemorrhage protocol, it was decided to apply these readily-available labels of haemorrhage to service two clinically relevant definitions as *appropriate MHPA activation*.

3.4.8.1 *Definition one; Major Haemorrhage:* MHPA deemed appropriate if patients receive more than or equal to four units of packed red blood cells within 24 hours of injury.

Definition one allows an assessment of performance in patients with less severe bleeding. In Chapter 2 clinicians reported that not all decisions were difficult but patients with less severe bleeding represented a difficult group of patients. These patients therefore represent a population of interest to assess the potential benefit of decision support.

3.4.8.2 *Definition two; Massive Haemorrhage:* MHPA deemed appropriate if patients received more than or equal to ten units of packed red blood cells within 24 hours of injury

OR

Patients who died due to uncontrolled haemorrhage

Despite its limitations this massive haemorrhage definition is the most commonly used definition of massive haemorrhage (**Table 3**). Using this same standard facilitates comparison with other studies.

3.4.9 Statistical methods

Categorical variables are expressed as frequency and percentage and analysed with Fisher's exact test. The normality of continuous variables was assessed using Q-Q plots and the Shapiro-Wilk test. Non-normally distributed data are expressed as the median and interquartile range (IQR) and were analysed with the Mann-Whitney U-test. Sensitivity and specificity were calculated in the standard manner. Confidence intervals for sensitivity and specificity are calculated with the Clopper-Pearson method. Accuracy was calculated as the overall probability that a patient will be correctly classified. McNemar's test was used to compare the sensitivity and specificity between the models and clinical judgement.

Youden's index was calculated as $(sensitivity + specificiy - 1)$. The index has a range of 0 to 1 and gives equal weight to false positive and false negative values. A perfect test has no false positives or false negatives and has a value of 1. The index can be defined for each point along a ROC curve. The maximum value of Youden's index was used as a criterion to select the optimum cut-off point in relation to both sensitivity and specificity.[388] To compare model performance of correlated data, DeLong's non-parametric comparison was used.[389]

Statistical analysis and figure development were performed using SPSS 26 (IBM, Armonk, New York, USA), Prism 8 (GraphPad Software, San Diego, California, USA) and R statistical software (R Foundation for Statistical Computing; www.r-project.org version 3.6.0).

3.5 Results

During the study period, 858 eligible patients were enrolled in ACIT. In the first 24 hours after injury 161/858 (19%) received at least 4units PRBCs, and 49/858 (6%) received 10units or more. The cohort was 82% male, 79% sustained a blunt MOI and had a median injury severity score of 13 (IQR 5-25). Five patients had a cause of death attributed to uncontrolled haemorrhage and received less than 10 units of PRBCs (**Table 12**).

Table 12: Patient characteristics, outcomes and model scores

	Total population	0-3 units PRBCs	4-9 units PRBCs	≥10 units PRBCs
n (%)	858	697 (81)	112 (13)	49 (6)
Patient Characteristics				
Age, years	36 (24-51)	35 (24-49)	47 (30-63)	38 (24-55)
Male gender	703 (82)	580 (83)	91 (78)	32 (73)
Blunt	679 (79)	545 (78)	92 (82)	42 (86)
SBP, mmHg[1, 2]	132 (114-149)	135 (120-151)	101 (80-130)[3]	91 (66-113)[3]
Lactate, mmol/L[1, 4]	2.2 (1.3-3.5)	1.9 (1.2-2.9)	3.4 (2.3-6.0)	7.5 (4.7-12.6)
Injury Severity Score[5]	13 (5-25)	10 (4-20)	33 (22-38)	30 (25-45)
In hospital mortality	83 (10)	28 (4)	27 (24)	28 (57)
PRBCs / 24 hrs, units	0 (0-1)	0 (0-0)	5 (4-7)	14 (11-22)
Native clinical judgement				
MHPA in either PH or ED	152 (18)	33 (5)	77 (69)	42 (86)
MHPA in PH phase only	111 (13)	26 (4)	51 (46)	34 (69)
Model scores				
ED TIC BN probability[6]	1 (0-5)	1 (0-2)	19 (6-38)	58 (29-73)
ABC score[7]	0 (0-1)	0 (0-1)[8]	1 (1-2)[8]	2 (1-2)[8]

[1] At the time of ED arrival

Missing data [2] 38 missing SBP data, [3] More missing SBP data in the ≥10 PRBC group than 4-9 PRBC group (31% vs 11% p = 0.002), [4] 47 missing lactate data, [5] 42 missing ISS data, [6] 858, [7] 41 missing ABC score data, [8] More data is missing 0-3units (2%), 4-9units (11%) and ≥10 units (31%). By comparison to the lower unit group there is more missing data at each higher unit group: all $p < 0.001$)

ABC; Assessment of Blood Consumption, ED; Emergency Department, MHPA; Major Haemorrhage Protocol Activation, SBP; Systolic Blood Pressure, PRBC; Packed Red Cell Volume, PH; Pre-hospital

There was no missing information for clinical activation of the major haemorrhage protocol. The ED TIC BN was calculatable in all patients in the study. The ABC score could not be calculated in 41/858 (5%) of the patients due to a missing variable in one of the four elements of the score. The most commonly missing variable was the ED arrival SBP in 38/41 patients.

3.5.1 **The accuracy of native clinical judgment**

Clinicians activated the major haemorrhage protocol in 152/858 (18%) of patients. The majority of MHP activations (111/152 (73%) occurred pre-hospital. Performance metrics for appropriate MHPA are given in **Table 13** and **Table 14**.

For the first definition (four or more units of PRBCs) clinicians had a sensitivity of 119/161 (74%) and an over-triage rate (1-specificity) of 33/697 (5%). For the pre-hospital activations clinical judgement had a sensitivity of 53% and an over-triage rate of 4%.

For the second definition (10 or more units of PRBCs) clinical judgement had a sensitivity of 42/49 (86%) and an over-triage rate of 110/809 (14%). For the pre-hospital activations clinical judgement had a sensitivity of 69% and an over triage rate of 10%.

Table 13: Classification table of clinical judgement, the ED TIC BN and the ABC score for two definitions of appropriate MHPA

		≥4 PRBCs	<4 PRBCs	≥10PRBCs	<10 PRBCs
Clinicians					
	yes	119	33	42	110
	no	42	664	7	699
ED TIC BN					
	yes	121	59	47	118
	no	40	638	2	691
ABC Score[1]					
	yes	53	43	20	76
	no	81	640	14	707

[1]ABC scores do not add to 858 due to missing data

3.5.2 Comparison of the ED TIC BN to native clinical judgement

For the first definition of appropriate MHPA, at the default ED TIC BN operating threshold of 10%, there was no statistical difference between the ED TIC BN and native clinical judgement in sensitivity (74% vs 75%, p = 0.87) however, clinicians were more specific (95% vs 92%, p = 0.002). For the second definition, despite the improved sensitivity of the model this did not reach statistical significance (86% vs 96% p = 0.13). At the second definition clinicians remained more specific than the model (86% vs 85%, p = 0.02).

Table 14: Diagnostic test results for two definitions of appropriate MHPA

	≥4 PRBCs			≥10PRBCs		
	Clinicians	**ED TIC BN[1]**	**ABC Score**	**Clinicians**	**ED TIC BN[1]**	**ABC Score[2]**
Sensitivity, %	74	84	40	86	96	59
Specificity, %	95	89	94	86	86	90
Positive Predictive Value, %	78	63	55	28	29	21
Negative Predictive Value, %	94	96	89	98	100	98
Accuracy, %	91	88	85	86	86	89

[1] ED TIC BN at optimised operating points using Youden's index for optimisation
[2]Data were only available for 34/49 patients

3.5.3 Optimisation of the ED TIC BN operating threshold

The ED TIC BN's output is a probability which can be dichotomised at any point between 0-100%. Adjustment of the operating threshold of the ED TIC BN changes the relationship between sensitivity and specificity. For the first definition of appropriate MHPA Youden's index was maximised at an operating threshold of 5.4%. At this operating point the ED TIC BN was more sensitive than clinicians (74% vs 84%, p = 0.01) and less specific (95% vs 89%, p < 0.001).

For the second definition of appropriate MHPA the same approach was used. At a threshold of 12.5% Youden's index was maximised. At this operating point the ED TIC

BN has a sensitivity of 96% and a specificity of 86%. This threshold has been labelled Operating Point 2 in **Figure 19.** Two further threshold values are used to illustrate the trade-off between sensitivity and specificity; the original threshold used in section **3.5.2** (10%, labelled Operating Point 1), and a threshold value selected with the same sensitivity as observed with clinical judgement (16%, labelled Operating Point 3). This third value was chosen to assess whether specificity could be improved relative to clinical judgement whilst maintaining sensitivity.

At operating point 2 there is no statistical difference between clinical judgement and the ED TIC BN in sensitivity (86% vs 96%, p = 0.13). At this operating point clinical judgement is no longer more specific than the model (86% vs 86%, p = 0.68). The overall diagnostic accuracy of the ED TIC BN at operating point 2 is 86%. Moving the threshold value to operating point 3 did not change these relationships with sensitivity (86% vs 88%, p = 0.13) or specificity (86% vs 89%, p = 0.09).

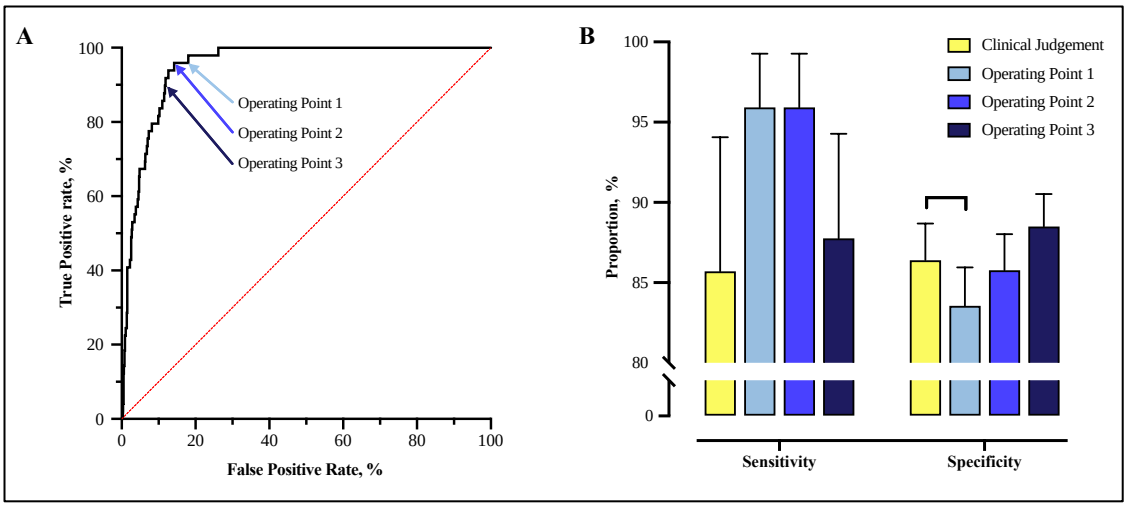

Figure 19: Changes in operating threshold of the ED TIC BN in patients receiving ≥10unit PRBCs **(A)** Receiver operating characteristic curve of the ED TIC BN. Arrows point to the position on the ROC curve at which the test is dichotomised. Operating points (OP) 1, 2 and 3 are at ED TIC BN probability thresholds of 10, 12.5% and 16% respectively. OP 1 is the original operating point used in model derivation, OP 2 the maximal point of Youden's index and OP 3 is fixed on the same sensitivity as observed clinical judgement. **(B)** Diagnostic accuracy of clinical judgement and the ED TIC BN at three operating points. Clinical judgement and OP 1 are statistically different. No other comparisons between clinical judgement and the ED TIC BN are significantly different. The Bars denote 95% CI.

3.5.4 Comparison of clinical and model errors

The model and clinicians produced the same classification in 783/858 (91%) and 741/858 (86%) of patients in definition 1 and 2 of appropriate MPHA respectively. At the ≥4unit PRBCs definition, clinicians did not activate the MHP in 42 patients that received ≥4units of PRBCs (**Table 13**). In the clinician, false-negative patients, 19/42 were identified as MHP patients by the model at the original operating point. On the other hand, there were 40 false-negative patients classified by the ED TIC BN. In the ED TIC BN false negatives, 17/40 were correctly identified by the clinicians.

At the ≥10unit definition of appropriate MHPA, there were seven clinician false negative patients. All seven of these patients were correctly classified by the ED TIC BN at operating point 2. Conversely, there were two false-negative ED TIC BN patients and one of these patients was correctly identified by the clinicians.

3.5.5 The performance of the ABC score relative to native clinical judgement and the ED TIC BN

To compare the performance of the models across all potential operating points ROC curve analysis was undertaken (**Figure 20**). The ROC curves demonstrate that the ED TIC BN had a greater area under the curve at both definitions of appropriate MHPA; definition one ED TIC BN AUROC: 0.92 vs ABC score AUROC: 0.78, $p < 0.001$; definition two ED TIC BN AUROC: 0.95 vs ABC score AUROC: 0.84, $p = 0.001$. The dichotomous clinical judgement is demonstrated as a single point with error on these figures. The point sensitivity and specificity of clinical judgement appears to largely outperform the ED TIC BN at the first definition of appropriate MHPA (**Figure 20A)** but is less sensitive and specific than the ED TIC BN at the second definition of MHPA (**Figure 20B**).

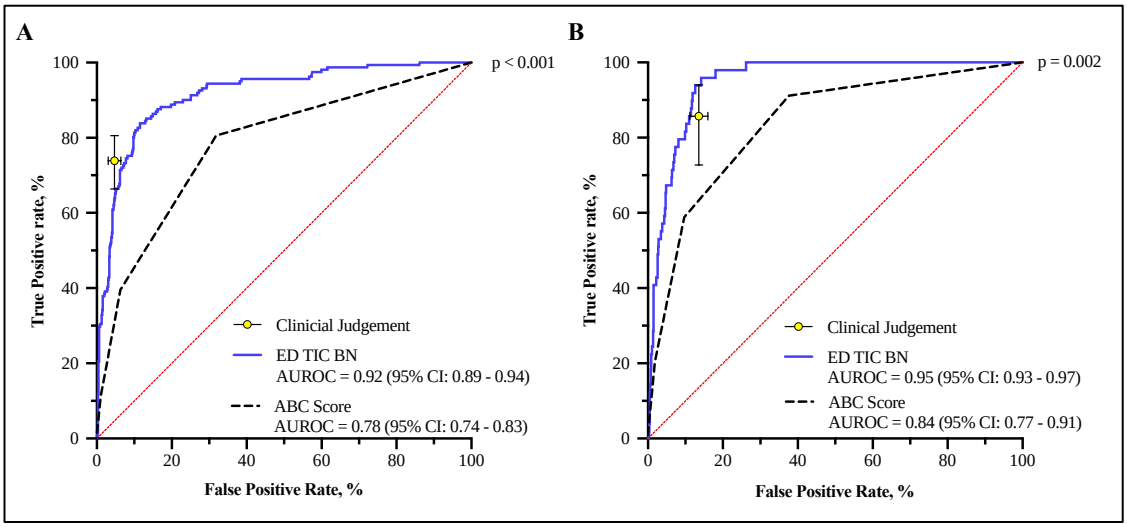

Figure 20: Comparison of the discrimination performance of the ED TIC BN and ABC Scores (A) Definition one of appropriate MHPA **(B)** Definition two of appropriate MHPA. Blue and black dashed lines represent the ROC curves for the ED TIC BN and ABC score models. The yellow circle represents the sensitivity and specificity of clinical judgement. Bars denote 95% CI.

3.6 Discussion

3.6.1 Key findings

As there is no agreed definition of 'appropriateness' of MHPA, available surrogate markers of haemorrhage were used. Using transfusion volume as a proxy for therapeutic need this study described the performance of clinical judgement and two models, in predicting the likelihood of appropriate MHPA. Clinical judgement exhibited good specificity with moderate sensitivity. Three quarters of MHPA occurred pre-hospital and pre-hospital activations were also characterised by high specificity and moderate sensitivity. In most circumstances the study did not identify a difference in performance between clinicians and the ED TIC BN. However, the ED TIC BN was shown to be more accurate than the ABC score. These encouraging findings represent the first assessment of the ability of the ED TIC BN to predict haemorrhage.

Compared to clinical judgement the ED TIC BN performed well in the prediction of clinically relevant haemorrhage. There were minimal differences between clinical

judgement and the ED TIC BN in haemorrhage prediction. The main difference is at the lower severity of appropriate MHPA where clinical judgement is marginally more specific than the ED TIC BN. At the higher severity definition clinical judgement and the ED TIC BN were not different.

The ED TIC BN produced a probabilistic output that was highly discriminant between different patient groups. This ability of the ED TIC BN model to accurately discriminate provides users with an ability to optimise performance for different patient groups.

3.6.2 **Implication of findings**

To consider the implications of the study's findings the generalisability of the included patient population will first be assessed. In a multicentre review of mortality in trauma haemorrhage 484 patients from 22 UK hospitals were included.[387] Between the present study and this national review, key metrics such as the 4 and 10 unit mortality were similar (4unit: 27% vs 24% and 10unit: 38% vs 53%). The number of units of PRBCs in the first 24 hours (4unit: 7units vs 5units and 10unit: 15units vs 14units), relative proportion of massive haemorrhage to major haemorrhage patients (3:1 vs 2.3:1) and arrival physiological observations are similar. Importantly, in both studies, the number of units of blood transfusion was proportional to the rate of in-hospital mortality. The similarity of patient groups between the two studies increases confidence of the representativeness of the patient sample in this study.

The generalisability of clinicians' expertise should also be considered. This study was conducted with clinicians at the Royal London Hospital. It is well documented that trauma patient volume varies between MTCs in the UK.[13] It is not known whether increased patient volume leads to improved clinical judgement. If the performance of clinical judgement is reduced outside the busiest trauma centres, decision support may be more impactful than demonstrated in the present study.

3.6.2.1 The performance of Clinical Judgement, the ED TIC BN and the ABC score

By comparison to other studies the clinicians in this present study perform extremely well. Two previous studies compare a CDST to clinicians in bleeding patients. The first study by Pommerening et al. is a prospective analysis of clinicians.[99] In 10 US level one trauma centres surgeons had a 66% sensitivity and 64% specificity for patients requiring ≥10 units PRBCs. Clinicians in this present study have an 86% sensitivity and specificity for the same end point. One important difference is the entry criteria into the two studies. In Pommerening's study all included patients had at least one unit of PRBCs whereas this present study is a study of all several injured trauma patients; many of whom receive no blood at all. It appears harder to separate patients that have some bleeding than patients with no bleeding.

Another possible explanation for the difference between studies is the time when the clinical judgement was made. In the Pommerening study clinicians were asked to make their decision at 10 minutes. In this present study clinicians could make their decision at any time while in the ED. It is likely that the addition of time and diagnostic adjuncts, like cross sectional imaging, improves clinical decision making.

The Pommerening study also examines the performance of common major haemorrhage models. The study suffers from missing data and was only able to calculate model scores in 486/1245 (39%) of the eligible population. Using a comparison of ROC curves the study demonstrates no difference between clinical judgement and the ABC and McLaughlin scores (AUROC 0.62, 0.63 and 0.66 respectively). The study reports that clinical judgement was less accurate than the TASH score (0.62 vs 0.72, $p = 0.01$). For comparison, the AUROC for clinical judgement, the ED TIC BN and the ABC score in the present study are 0.86, 0.95 and 0.84. Thus, it appears both clinicians and the models perform better in this current study.

The second study to compare clinicians and a major haemorrhage model, reports the positive predictive value (PPV) for patients receiving >5units PRBCs. In this retrospective study, clinical judgement and the ABC score have 73% and 34% PPV respectively. Unfortunately, sensitivity and specificity were not reported and cannot be

calculated from the summary data presented. The authors conclude that ABC driven activation of the MHP would result in earlier activation at a cost of a seven-fold higher blood product wastage.[233] Differences in study design and haemorrhage prevalence make it difficult to compare performance between studies. Notably, the general relationship between the PPV of clinical judgement and the ABC score is the same in both studies, i.e. clinical judgement had a higher PPV.

What this second study adds is an understanding of the influence of a CDST on the time to MHPA. The study suggests, ABC activated MHP would reduce time to activation by 35 minutes compared with physician judgment.[233] An accurate model that was able to reduce the time to MHPA would have considerable clinical utility. It has been shown that delays to the arrival of blood products is associated with decreased survival.[94] From recent direct observation of the ED TIC BN the model provides its most accurate prediction at 15 minutes from the time of patient arrival in the ED (see **1.9.5**). The median time to MPHA activation in the Motameni study was 37 minutes (IQR 26-52 minutes). Unfortunately, in this present study, the time of MHPA was not documented. If MHPA in our setting is similar, to the Motameni study the majority of ED TIC BN activated MHP would be faster than clinician activation.

3.6.2.2 The potential of the model to support clinical decisions

Having established that clinical judgement could be improved with respect to activation of major haemorrhage protocols, the next step is to evaluate the value that the ED TIC BN model might add to the decision-making process. As the ED TIC BN's accuracy is less than 100%, analysis of human and model errors maybe able to identify systematic errors to guide the use of the model. In this study clinical judgement and the ED TIC BN categorised patients in the same direction in 86-91% of cases. In cases of disagreement neither the model nor the clinician were more likely to be 'right'.

The Youden index derived "optimisation" of the operating point of the ED TIC BN created CDSTs that were more sensitive and less specific than clinical judgement. It is important to note that Youden's index gives equal weight to sensitivity and specificity. In clinical practice the trade-off between false negatives and false positives may not be

given equal weight. In a clinical problem like major bleeding the majority of the patients do not have the disease; there is a class imbalance. Increasing the sensitivity of a test and reducing the number of false negatives may benefit some patients but comes at a cost of over-triage. With the class imbalance seen in this problem, small decreases in specificity proportions lead to greater raw numbers of over-triaged patients. Unnecessary activation of the MHP will have costs in terms of wasted resources, potential harm to patients that won't derive therapeutic benefit and disruption of the normal activities within a major trauma centre.

The potential to improve clinical decision making is not only dependent on the value ascribed to each of the sensitivity and specificity issues but also the way clinicians interact with the model. It is not apparent how clinicians might behave in cases of disagreement between clinical judgement and the model. In a series of experiments conducted by Mossadegh, she identified that the direction of change (e.g. from a position of low risk to high risk) and the seniority of the clinician influence the likelihood that the clinician will change their decision given the model's output (see **1.9.5**). [327, 341]

3.6.2.3 Military and austere value of the ED TIC BN

The ED TIC BN has clinical value in the austere or military setting. First, it may enable less experienced clinicians to perform with an accuracy that is not significantly different to expert clinicians. Second, by changing the operating threshold for the ED TIC BN, a trauma system would be able to prioritise either the test's sensitivity or specificity. For example, a need to prioritise specificity may be desirable during resources shortages. In a military context adjusting the operating threshold creates a decision rule which is adaptable in real time. This provides flexibility for senior policymakers while ensuring tactical level decisions remain unburdened by unnecessary complexity.

3.6.2.4 ED TIC BN vs ABC Score

The ED TIC BN has demonstrated superior predictive performance in this study than the ABC score. The ED TIC BN was not trained to predict major haemorrhage but is able to effectively risk stratify bleeding patients. The simplicity of the ABC score

adversely affects its predictive performance and it offers little opportunity to augment the clinical judgement demonstrated in this study.

3.6.3 **Strengths and Limitations**

3.6.3.1 Strengths

By using data collected from a prospective observational study the quality of the data is high. In addition, this study's retrospective design provides a large number of patients to be analysed. This allows rare outcomes like massive haemorrhage to be studied in the sample population. High quality and volume of data has demonstrated the uncertainty of MHPA in patients treated at the Royal London Hospital.

The study uses two clinically relevant definitions of major haemorrhage. As noted in previous research[387] the volume of blood transfusion was positively correlated with mortality. In this study patients that received 0-3, 4-9 or 10 or more PRBC units correlated with in-hospital mortalities of 4%, 24% and 57% respectively. It is vitally important that a decision support tool is able to discriminate, and risk stratify patients in this middle group. Using only the historical definition of Massive Transfusion, here definition two, would ignore a group with a 1 in 4 risk of death.

The definitions of appropriate MHPA reflect the delivery of blood and blood products within the hospital's MHP. The RLH's MHP includes the immediate delivery of four units of both PRBCs and Fresh Frozen Plasma (FFP) and subsequently if a second pack is required 6 PRBCs, 6 FFP, two cryoprecipitate and one pool of platelets. As such, a patient that receives all of the PRBCs in the first pack has been classified as a major haemorrhage patient, and a patient that receives all of the PRBCs in both packs is classified as a massive haemorrhage patient.

3.6.3.2 Limitations

This study has limitations that can be considered in three domains; the study design, the clinical application, and the modelling execution.

The study's retrospective design predisposes it to common biases. For example, there are more missing ABC scores amongst patients with the highest mortality. This may represent a systematic selection bias and casts questions over the performance of the ABC model in the highest risk group. Another problem is the lack of important variables such as the time of MHPA. This variable would be particularly helpful to understand whether the ED TIC BN model could be used in a timely fashion to assist decision making.

There are limitations associated with the patient population in the study. The data from this study all comes from a single centre and was collected several years ago. The ACIT study population is also a subset of trauma patients. In a study from the same institution 4% of all trauma patients activated the MHP.[170] In the ACIT patient population, the rate of MHPA is 18%. This suggests that the ACIT population is a more severely injured sample of patients compared to the general trauma population. These factors affect the generality of the study's findings to a broader trauma population.

Despite these limitations, the retrospective design has provided a large quantity of high-quality data. The higher severity of trauma patients has provided more high-risk patients which has allowed the study to assess the performance in the patients of most interest. The data has allowed the study to effectively assess the hypotheses within these common limitations of retrospective research.

The next area of limitation is the clinical validity of the study. This concerns the appropriateness of the definitions chosen. An appropriate definition of critical bleeding early after injury remains an area of active research with no currently accepted consensus.[390] The definition of massive transfusion of 10 units of PRBCs in 24 hours has been criticised as arbitrary, subject to survivor bias, and neglectful of transfusion rate early after injury.[391] To address these concerns, this study used two definitions of appropriate MHPA. These definitions have provided a pragmatic sensitivity analysis based on previously accepted definitions.

Newer research tools are gaining traction that are designed to identify bleeding patients at risk. Tools such as the Critical Administration Threshold (CAT)[391] and the

Resuscitation Intensity (RI)[392] have been proposed to define the population at risk more appropriately. These newer methods simultaneously account for modern balanced transfusion practice.[390] Recognition of the value of these newer methods to identify the high risk patient groups is developing. The benefit of this changing approach has been demonstrated in the recent re-interpretation of a high profile trial.[164] By changing how high risk bleeding patients were defined in the trial the authors were able to improve the clinical validity of the study results. Using a technique such as the CAT or RI to define MHPA may enhance the analysis of future research like the present study.

Defining appropriate activation of an MHP is complex as it is a multifaceted judgement. It requires the accurate identification of high-risk patients, as just discussed. But it also requires a multidimensional decision about what constitutes appropriate (or justified) use of the MHP. These dimensions include impacts on the patient but also the health care system. For example, what is appropriate in one healthcare setting may not be appropriate in another. If resources are limited, then wastage of blood and blood products maybe given more weight when the patient benefit is less certain. This study is not able to address this complex utility decision.

Another area of clinical validity to consider is the range of processes an MHP includes. As discussed in section **1.4.5.2** a major *haemorrhage* protocol (MHP) is not purely a massive *transfusion* protocol. Other elements of an MHP include rapid assembly of senior decision-makers and access to haemorrhage control strategies. It is possible that a clinician activated the MHP to gain rapid access to expertise or facilities, without the need for large volumes of blood transfusion. For example, with an isolated extremity injury temporary haemorrhage control may be achieved with a tourniquet. The physician may still trigger the MHP for rapid surgical repair. However, this patient may not need large volumes of blood transfusion once haemorrhage control is achieved. In this study, this isolated extremity injury scenario will make clinicians' decision accuracy appear poorer. This is because accuracy of appropriate MHPA has been measured by a transfusion volume. This definition of appropriate MHPA does not consider other factors such as the speed of decision making or patient transport to the operating room. This overall impact of this effect on the study is unknown.

The third area of limitation pertains to the modelling approach. The ED TIC BN model has been used to predict major haemorrhage. The model was not developed for this outcome and has not been trained to make this prediction. While TIC and major haemorrhage are intricately linked a model trained specifically to predict major haemorrhage may perform more accurately. Equally, the volume of blood a patient requires, and the volume they receive are not necessarily the same. It is difficult to quantify "unnecessary" transfusion.

The ideal modelling approach would quantify the amount of blood the patient has lost as a continuous variable. This outcome is preferred over an approach which models a clinical decision. This is because the decision of how much blood to give a patient will contain error and practice may evolve over time (see **2.5.2**). An approach that models how much blood the patient has lost, therefore models a physical property of the patient. This theoretically superior modelling approach requires an outcome (blood loss) that is much harder to measure.

The last modelling limitation concerns the study data. In this study 519/858 (60%) of the patients were also used to developed the ED TIC BN. The effect of using this development data to assess model performance is an over-estimation of the ED TIC BN's performance. A more appropriate method would only use patients from a population other than the model development population. Reassuringly, there was very little difference in the performance of the ED TIC BN model between the development and validation cohorts (see **1.9.3**).

3.6.4 **What remains unknown**

This study is unable to comprehensively explain how the tool will influence decision making in the real world. Comparing clinicians and the ED TIC BN in parallel as described here, does not address the interaction of the clinician with the model. The five key steps in blue boxes in **Figure 18** have not been addressed in this study.

Evidence suggests with accurate models clinical prediction can be improved when clinicians and a model are combined.[393] However, only an impact analysis can determine whether the use of the model is better than usual care. For example, it is not obvious what the clinician will do in cases were they disagree with the model's classification. Previous studies have identified complex interactions between clinicians and a CDST. In a study of a CDST for chest pain triage in an emergency department, clinicians and the CDST disagreed in 26% of patients. Interestingly, in cases where the clinicians disagreed or did not use the CDST, physicians' decisions were less efficient and less safe. The disagreements commonly arose when physicians overruled the CDST's recommendations.[218]

Future studies on the ED TIC BN should assess the interaction between the clinicians and the CDST. Assessment of the key steps in **Figure 18** is necessary to understand the clinical utility of the model. The resultant accuracy of supported decisions is dependent on understanding the human-CDST interaction. Key to understanding the human-CDST interaction and resultant impact of the model on clinical practice is ensuring the model is assessed in the intended population. In this study it was possible to quantify the performance of clinicians in the pre-hospital environment (**Table 12)** but not the models. To understand the impact of pre-hospital decision support studies should be conducted in the pre-hospital environment.

3.7 Conclusion

This chapter has quantified the predictive accuracy of native clinical judgement and a BN in decisions about major haemorrhage protocol activation. The study has several important findings. First, it has provided evidence of a clinical problem that stands to benefit from decision support. Second, it has applied the ED TIC BN to decisions about blood transfusion and assessed its performance. Until this point there had not been a demonstration of the clinical utility of TIC prediction. Third, the study demonstrates that expert clinical judgement and the ED TIC BN model are broadly similar, especially at the more severe definition of appropriate MHPA. The chapter has moved a step closer towards understanding the ED TIC BN's clinical impact.

The study noted 73% of clinical MHPAs were made pre-hospital. Native clinical judgement pre-hospital also exhibited only moderate sensitivity and good specificity. As the majority of decisions around major bleeding are made pre-hospital, and earlier intervention maybe related to better patient outcomes, the value of BN decision support pre-hospital warrants further analysis. Future, prospective pre-hospital analysis will assess whether a BN could usefully augment decision making pre-hospital and reduce the intrinsic biases of retrospective research.

The next chapter takes the existing ED TIC BN and assesses how it might perform when modified for pre-hospital use.

CHAPTER 4

PUSHING FORWARDS: FROM EMERGENCY DEPARTMENT TO PRE-HOSPITAL DECISION SUPPORT

4.1 Scope of the chapter

This chapter examines the feasibility of adapting the ED TIC BN model for TIC prediction in the pre-hospital environment. From the previous chapter, it was apparent that the majority of decisions for activation of the major haemorrhage protocol were made pre-hospital. These decisions contain errors and decision support may provide an opportunity to improve patient outcomes. Supporting decisions pre-hospital requires models that are designed to function in this environment. This chapter, first, assesses the variables available in the pre-hospital setting; second, adapts the *emergency department* TIC BN to a *pre-hospital* TIC BN; third, assesses the plausible performance of the adapted PH TIC BN using an existing trauma patient registry and fourth compares the performance of the PH TIC BN to two previously published PH TIC models.

4.2 Introduction

The modern management of major haemorrhage and TIC emphases early intervention (see **1.4.5**). An effective way to shorten the time between injury and intervention is to initiate treatment before a patient arrives in hospital.[36, 185] A barrier to early intervention arises when high risk patients are difficult to identify. Chapter 2 demonstrated that decisions around major haemorrhage and TIC are made difficult by clinical uncertainty and imperfect decision evaluation. These conditions reduce the accuracy of native clinical judgement as demonstrated in Chapter 3. Errors in critical decisions, such as delayed recognition of a high-risk patient or failure to provide a beneficial therapy, can have profound consequences. Working on the premise that faster access to appropriate treatment improves outcomes in trauma (**see 1.4.3**), this thesis aims to support decision making at the earliest opportunity after injury. Clinical Decision Support Tools (CDST) offer an opportunity to overcome barriers to early intervention by reducing uncertainty and improving decision making.

The ED TIC BN described in section **1.9**, was developed and validated for use as a CDST *in*-hospital in the Emergency Department. The ED TIC BN has 14 input variables obtained during an ATLS primary survey (**Figure 12**). The relative lack of diagnostic aids pre-hospital and earlier post-injury physiology is likely to lead to a deterioration in the model's performance. It is not known whether the predictive accuracy of the ED TIC BN is sufficiently maintained when used with pre-hospital information to augment decision making. This chapter will examine the potential for transformation of the *ED* TIC BN to a *PH* TIC BN.

As discussed in Chapter 1, BNs have several valuable features, such as incorporation of domain knowledge and the ability to model complex non-linear variable relationships (**see 1.8.2.2**). These features may enable a BN model to outperform models which rely on traditional statistical techniques. Once a PH TIC BN model has been constructed this chapter will assess the performance of the PH TIC BN compared to published PH TIC models.

4.3 Aims and Hypotheses

4.3.1 Aims

4.3.1.1 *Aim 1:* Identify the set of variables available for pre-hospital model use

4.3.1.2 *Aim 2:* Assess the performance of the adapted PH TIC BN

4.3.1.3 *Aim 3:* Compare the adapted PH TIC BN to other published PH TIC models

4.3.2 Hypotheses

4.3.2.1 *Hypothesis 1:* Not all 14 variables for the ED TIC BN are available pre-hospital

4.3.2.2 *Hypothesis 2:* Performance of a PH TIC BN, using the same model structure and parameters as the ED TIC BN, is sufficiently maintained to warrant further clinical examination

4.3.2.3 *Hypothesis 3:* Performance of the resultant PH TIC BN is superior to other published models

4.4 Methods

This chapter follows the guidance of the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.[394] The TRIPOD statement aims to improve the reporting of development, validation and updating studies of prediction models.

4.4.1 Availability of input variables for a PH TIC BN

The variables available for a PH model were established by observation of care at LAA and AAKSS. Pre-hospital blood gas analysis is not currently performed by either Air Ambulance service. During this research, clinicians at AAKSS judged the feasibility of performing blood gas analysis between the point of injury and arrival in hospital. A

portable blood gas analyser (epoc® Blood Analysis System (Siemens, Camberley, UK)) was used during this assessment.

Additionally, representatives of Air Ambulance services in the UK were asked in May 2020 whether their service had access to blood gas analysis. Respondents were members of an online group of currently practicing pre-hospital care clinicians.

4.4.2 Defining the PH TIC BN

Following observation of the potentially available model inputs, a new PH TIC BN was defined. The PH TIC BN model retained the same structure and parameters as the ED TIC BN (see **1.9)**. Only the input variables changed between PH and ED models. The PH model's input variables were either retained, modified or removed from the original ED model. The decision to retain, modify or remove an input variable, depended on the pre-hospital availability of that variable.

4.4.3 Study design to assess PH TIC BN performance

To assess the performance of the PH TIC BN, the model's discrimination, calibration and accuracy is assessed in a cohort of ACIT study patients. These performance methods are discussed in **4.4.8**.

The ACIT database does not record the PH clinician's injury assessment and diagnosis. Instead, ACIT only contains the final injury diagnosis on hospital discharge. This presents a problem on how to validate the PH TIC BN model. The PH TIC BN uses inputs concerning the presence of Haemothorax, Abdominal Bleeding, Open Long Bone Fracture and Unstable Pelvic Fracture, diagnosis of which is supported by investigations that are not available to the pre-hospital clinician. Therefore, for the purposes of this study, assumptions that a particular injury input would have been amenable to accurate PH diagnosis had to be made. It was assumed that PH clinicians encountering an ACIT patient later logged as having unstable pelvic fracture and/or long bone fracture would been able to make these diagnoses accurately, as the clinical signs associated with these diagnoses are usually overt. For haemothorax and intra-abdominal bleeding, clinical

signs may be less obvious to the PH clinician and for that reason a higher threshold of documented injury severity was applied. As such, it was assumed that patients with chest or abdominal injuries later coded as Abbreviated Injury Scale (AIS)[386] ≥3, and whose free-text description of the injury was consistent with the presence of haemorrhage, would have had displayed sufficient clinical evidence of the presence of these injuries to make the diagnosis. Therefore, these input variables (Haemothorax and abdominal) have been assumed to be accurately diagnosed when AIS≥3 and have been entered into the model in this study.

To evaluate the potential for overfitting, the performance of the PH TIC BN was also assessed in a cohort of patients that were not used in the development of the original ED TIC BN model. This external validation cohort was the same group of patients previously used to externally validate the ED TIC BN.[338]

4.4.4 Study population

To assess model performance, the same cohort of patients from the ACIT study as used in **CHAPTER 3** was used again. A full description of the study setting, and patient characteristics are provided at **3.4**. The ACIT study prospectively collects multiple patient data points at several time points. These data points include patient demographics, mechanism of injury, injuries sustained, vital signs, treatment administered and outcome. Pre-hospital vital signs are collected in ACIT and were not used in the development or validation of the ED TIC BN.

4.4.5 Comparison of the PH TIC BN to other models

There are two published prediction models for PH TIC: the Coagulopathy of Severe Trauma (COAST) Score[257] (see **1.6.3.2**) and the Prediction of Acute Coagulopathy of Trauma (PACT) Score[261] (see **1.6.3.3**). The models differ in their modelling approach, complexity, inputs, outcome classification and predictive performance. (**Table 4** and **Table 15**).

Table 15: Input variables and characteristics of the pre-hospital Trauma Induced Coagulopathy prediction models

Input variables	COAST		PACT		PH TIC BN	
Demographics						
			Age	Decade		
Vital Signs						
SBP, mmHg	<100 or <90		HR/SBP ≥1	Yes/no	SBP	Cont.
Temperature, ºC	<35 or <32		GCS	3-15	HR	Cont.
					Temperature, ºC	Cont.
					GCS	3-15
Mechanism of injury						
					Energy	High / low
Causes vehicular entrapment	Yes/no		Non-vehicular	Yes/no	MOI	Blunt / pen.
Suspected injuries						
Abdominal or pelvic injury	Yes/no				Long bone #	Yes/no
					Unstable pelvic #	Yes/no
					Haemothorax	Yes/no
Investigations					Abdominal bleeding	Yes/no
					Lactate	Cont.
Treatments received						
Chest decompression	Yes/no		Intubation	Yes/no	Pre-hospital Crystalloid	Cont.
			CPR	Yes/no		
Model characteristics						
Number of variables	5		6		12	
Model design	Simple score		Logistic regression		Bayesian Network	
Scoring	Integers (0-7)		Integers (0-400)		Probability (0-1)	
Model application	Paper form		Online calculator		Online interface	

SBP, Systolic blood pressure; GCS, Glasgow Coma Scale; HR, Heart Rate; CPR, Cardiopulmonary resuscitation; MOI, Mechanism of injury; #, Fracture; Cont., Continuous; Pen., Penetrating.

The COAST and PACT models were applied to patient data in this study using the originally published model descriptions. The COAST and PACT scores can be readily applied to the data. The PH TIC BN requires either the original rendering of the BN in AgenaRisk software[323] or an online model calculator available at www.traumamodels.com.

4.4.6 Outcome classification

In common with much of the TIC literature, the three models differ in their classification of TIC (**Table 30**). The COAST score classified TIC as INR > 1.5 OR aPTT > 60 seconds. The PACT score classified TIC as an INR > 1.5. Models trained to predict diagnostic test results, rather than the underlying pathological condition, compound errors. To mitigate against additional error, the ED TIC BN classified patients as coagulopathic using a threefold approach; coagulopathic patients had an INR >1.2[395] AND were identified as coagulopathic by an expectation maximisation (EM) clustering algorithm.[396] The EM algorithm grouped patients into two clusters using the patient's clinical, laboratory, and thromboelastometry profiles.[335] If the INR and EM clustering methods produced conflicting results, a third step was performed. In the deciding step, an expert review was performed using all the available patient information.[338]

In this study, a patient's coagulopathy status is classified using the threefold method established by Perkins and Yet.[338]

4.4.7 Handling and substitution of input data

Simple scores and logistic regression models require a complete set of input variables to perform in the manner they were developed and validated. Bayesian Networks are not constrained to complete case analysis and will provide outcome calculations without a full set of input variables. Therefore, the COAST score and PACT score required complete-case analysis to calculate their predictive performance. An *a priori* decision was made to calculate the prognostic performance of the COAST and PACT models under two conditions. First, a complete-case analysis was performed. Second, model

performance was re-calculated without the most common missing variable. This secondary analysis was performed to assess for evidence of selection bias resulting in poor performance due to missing data. A final assessment of predictive performance was compared for patients that had a complete set of input variables for the PH TIC BN. Missing data imputation was not performed for any model.

The ACIT database does not contain all of the variables required for the COAST, PACT or PH TIC BN scores. These missing variables were pragmatically approximated according to the rules in

Table 16**.**

4.4.8 **Statistical methods**

The normality of continuous variables was assessed using Q-Q plots and the Shapiro-Wilk test. Numerical data are reported as median (IQR) and categorical data as frequency (n) and percentage (%). Statistical significance was set as a 2-tailed P value of <0.05.

4.4.8.1 Model performance: calibration

Predictive performance was assessed in terms of *calibration* and *discrimination*. *Calibration* refers to how closely the predicted risk of TIC agrees with the observed TIC risk. A calibration plot was generated by comparing the average observed frequencies (y-axis) to the average predicted probability (x-axis) of the outcome. This was assessed for each decile of predicted risk. A smoothed curve was fitted to the data to assist with visualisation of the relationship using a regression technique (Locally estimated scatterplot smoothing; LOESS).[397] Using this method, perfect predictions are plotted on the ideal line, described with an intercept of 0 and a slope of 1. The intercept relates to calibration-in-the-large and compares the mean of all predicted risk with the mean observed risk. The slope provides information on whether the model under- or over-estimates risk across the range of model predictions. Hosmer-Lemeshow tests were not conducted as calibration plots are preferred.

Table 16: Rules of approximation for variables from ACIT database to pre-hospital TIC prediction models

Pre-hospital variable in prediction model	Available in ACIT data	Method of approximation
COAST		
SBP, mmHg	✓	
Temperature, ºC	x	Substituted ED arrival temperature
Vehicular entrapment	x	Assumed when time from injury to ED >1.5hrs and appropriate mechanism of injury
Abdominal or pelvic injury	x	Assumed if AIS abdomen or pelvic injury ≥3[1]
Chest decompression	x	Assumed if diagnosis of tension pneumothorax[1]
PACT		
HR/SBP ≥1	✓	
Age	✓	
GCS	✓	
MOI non-vehicular	✓	
CPR	x	Assumed if PH or ED observations demonstrate cardiac arrest
Intubation	✓	
PH TIC BN		
SBP	✓	
HR	✓	
GCS	✓	
Energy	✓	
MOI	✓	
Long bone #	x	Assumed if long bone fracture coded[1]
Unstable pelvic #	x	Assumed if unstable pelvic fracture coded[1]
Haemothorax	x	Assumed if AIS chest ≥3 and haemothorax coded[1]
Abdominal bleeding	x	Assumed if AIS abdomen ≥3 and injury consistent with bleeding coded[1]
Lactate	x	Substituted first in-hospital lactate
Pre-hospital crystalloid	✓	

[1]For the purposes of model calculations injuries were assumed to be apparent to pre-hospital clinicians if the injury was coded at the final diagnosis.
SBP, systolic blood pressure; GCS, Glasgow Coma Scale; HR, heart rate; CPR, cardiopulmonary resuscitation; MOI, mechanism of injury; #, fracture

4.4.8.2 Model performance: discrimination

Discrimination is the ability of the model to distinguish between patients who do and do not develop TIC. Discrimination can be visually assessed from a stratified histogram of predictions. This histogram was implemented as a *rug plot* nested within the calibration plot. Discrimination is calculated as the concordance statistic (c-statistic) which for a binary outcome is the same as the AUROC. The AUROC was calculated using Hanley's method.[398] A c-statistic of 0.5 represents chance, and a value of 1 is perfect discrimination. AUROCs were reported with 95% Confidence Intervals. Correlated data was present when ROC curves were generated from two or more models applied to data from the same individuals. To compare model performance of correlated data, DeLong's non-parametric comparison was used.[389] Hanley's method[398] was used for model comparisons of independent data.

4.4.8.3 Model performance: overall measures - Brier scores

In addition to calibration and discrimination, measures of overall model performance were used. The Brier score (BS) quantifies the overall accuracy of a probability model by taking an average of the errors of prediction from each patient.[399] The BS has a range of 0 (no error between prediction and outcome) and 1 (worst possible model). The Brier skill score (BSS) is a modification of the BS which divides the BS by the average probability of the event occurring. The BSS gives the improvement in using the model compared to predicting the average outcome. The BSS has a range from - ∞ to 1. Negative numbers demonstrate a worse prediction than the average probability, and a score of 1 is the perfect model. The 95% confidence interval of the BS and BSS are calculated using bootstrapping.

4.4.8.4 Model performance: overall measures - net benefit

Calibration and discrimination describe the predictive performance of a model, but do not provide any insight into the clinical consequences of miscalibration or imperfect discrimination. Decision Curve Analysis (DCA) is a novel method used to assess the potential population impact of a prediction model.[400] DCA estimates whether clinical

decisions from a model would do more good than harm i.e. the model's *net benefit.* Net benefit is calculated by weighing improved patient outcome against the harm to patients that cannot derive benefit. Net benefit requires a clinical judgement on the relative importance of benefits against harms.

Mathematically, DCA relates the number of true-positives (TP) predictions to the number of false-positive (FP) predictions at a given operating point or outcome risk threshold. The difference between TP and FP is then weighted by the factor that determines the value of a FP relative to a false negative (FN). This factor acts as an "exchange rate" between the errors of FP and FN and allows harm and benefit to exist on the same scale: *net benefit.*[401]

$$net\ benefit = \frac{TP}{N} - \frac{FP}{N} * \left(\frac{p_t}{1 - p_t}\right)$$

Where N is the total sample size and p_t is the operating point or threshold probability used to classify patients into risk categories for decision making. Notably benefit and harm are not included explicitly in this function, but are included implicitly in the rationale selection of p_t, reflecting the harm and benefit of intervention.[402]

For example, a PH TIC model may recommend patients should undergo damage control resuscitation (DCR) at a threshold of $p_t = 0.10$. Above this value all patients receive DCR. Using the net benefit equation allows a weighted calculation of harm to benefit at this threshold or exchange rate. Harm is the magnitude of the side effects of patients that unnecessarily receive DCR (FP) and benefit is the magnitude of improvement due to DCR for patients with TIC (TP). Net benefit can be calculated for any p_t from 0 to 1, to assess the relative harm or benefit at each given probability threshold.

In this study, the relationship of net benefit to p_t was plotted graphically as a decision curve at all values of p_t.[400] The decision curve plot can be viewed as a sensitivity analysis over a range of exchange rates between benefit and harm. With a given p_t the curve displays the net benefit of using the model at that threshold. The decision curve was used to plot the net benefit of the PH TIC BN and two default strategies: a strategy

to treat everyone and a separate strategy to treat no one.[400] The strategy or model with the highest net benefit at a given p_t offers the greatest clinical value at that threshold.[401]

4.4.8.5 Statistical computation and figure development

Statistical analysis and figure development were performed using SPSS 26 (IBM, Armonk, New York, USA), Prism 8 (GraphPad Software, San Diego, California, USA) and R statistical software (R Foundation for Statistical Computing; www.r-project.org version 3.6.0). "pROC" package version 1.16.2 was used to compare ROC curves.[403] Packages used to create the calibration plots and calculate Brier scores included "ggplot 2" version 3.3.0[404] and "rms" (Regression Modelling Strategies) version 5.1-4.[405] The decision curve analysis was created with the "rmda" (Risk Model Decision Analysis) version 1.6 package.[402] The BN was computed by AgenaRisk software (Agena, London, UK).[323]

4.5 Results

4.5.1 Available variables in the pre-hospital environment for a PH TIC BN

Of the original 14 input variables the PH TIC BN retained three, modified eight and removed three. The mechanism of injury, the energy associated with the injury and the PH crystalloid volume were retained. Of the eight modified input variables four were modified from ED to PH measurements: heart rate, systolic blood pressure, GCS and lactate.

Notably, the use of point of care (POC) lactate measurements was variable between Air Ambulance services.[114] One of the two observed sites in this study used POC lactate. However, the use of PH lactate measurement at this site was intermittent. During the period of observation, 15% of AAKSS patients had a POC lactate measured.[406]

The other four modified variables were the suspected injury variables. The ED TIC BN model utilized imaging to support the diagnosis of bleeding and fractures. These imaging modalities are not routinely available pre-hospital. For the PH model, variables relating to chest, abdominal, pelvic and long bone injuries have been modified to rely on clinical diagnosis alone.

Temperature was not routinely measured in trauma patients pre-hospital and was removed. Base deficit (BD) and pH were also removed. The decision to remove BD and pH followed clinical evaluation of a portable blood gas analyser at AAKSS. Clinicians reported the device took too long to operate, was prone to device errors, and was not easy to operate PH. In addition, clinicians at every Air Ambulance service in the UK responded to the survey. Only four of the 22 Air Ambulance services had access to blood gas analysis. The resultant PH TIC BN model consists of a maximum of 11 input variables (**Figure 21)**.

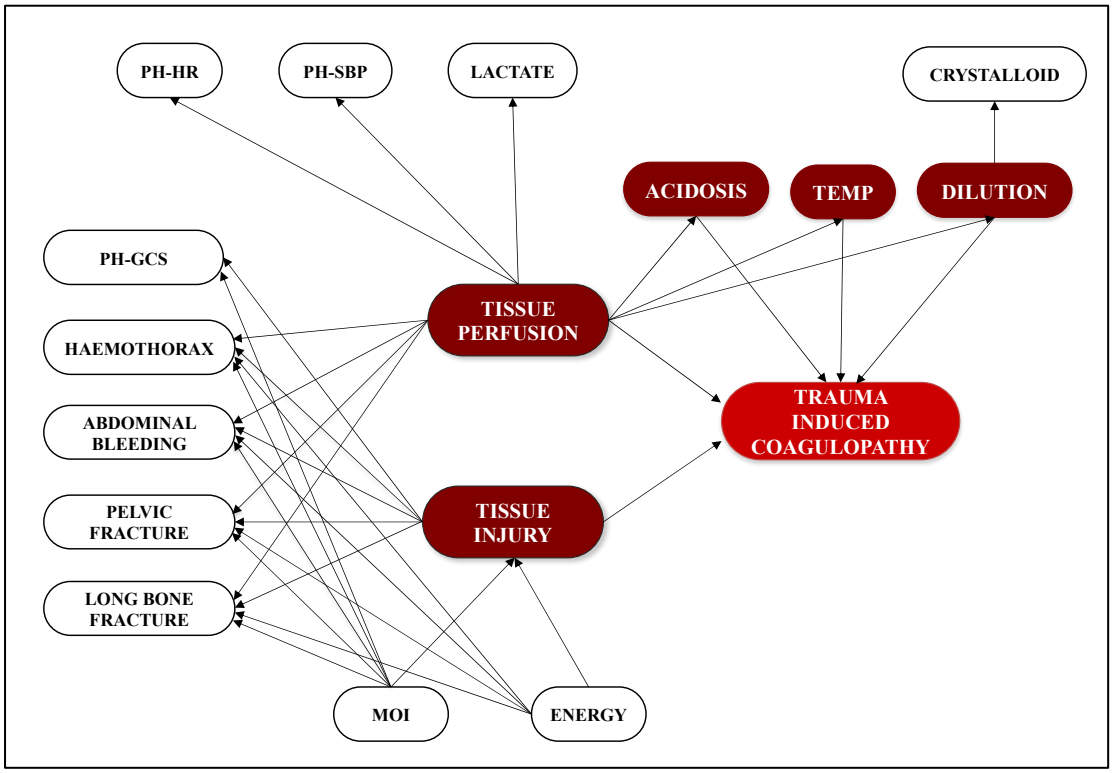

Figure 21: Directed acyclic graph of the pre-hospital trauma induced coagulopathy Bayesian Network. The bright red variable represents the predicted outcome. Dark red variables represent the five latent causal factors. The white input variables (evidence) represent predictors associated with the causal factors. Evidence variables have changed from **Figure 12** to variables available in the pre-hospital environment. PH, pre-hospital; GCS, Glasgow Coma Scale; HR, heart rate; MOI, mechanism of injury; SBP, systolic blood pressure; Temp, temperature.

4.5.2 Study Population and data availability

The study population is described in section **3.5, Table 12** and **Table 17.** The study included 858 patients, of which 82% were male, median ISS was 13 (IQR 5-25), and 92/858 (11%) had TIC. Despite the approximation methods, some variables had large proportions of missing data (**Table 17**).

Table 17: Missing data and summary of scoring variables for pre-hospital TIC prediction models

	Pre-hospital variable	Missing data	Cohort values
COAST			
	SBP mmHg, n (%)[1]	91 (11)	0, 634 (74) 1, 41 (5) 2, 92 (11)
	Temperature (°C[1])	346 (40)	0, 433 (50) 1, 78 (9) 2, 1 (0)
	Vehicular entrapment, n (%)	6 (1)	64 (8)
	Abdominal or pelvic injury, n (%)	26 (3)	147 (17)
	Chest decompression, n (%)	26 (3)	25 (3)
PACT			
	HR/SBP ≥1, n (%)	327 (38)	84 (10)
	PACT Age, median (IQR)	3 (0)	4 (2–5)
	PACT GCS, median (IQR)	41 (5)	0 (0, 3)
	MOI non-vehicular, n (%)	1 (0)	613 (71)
	CPR, n (%)	0	12 (1)
	Intubation, n (%)	1 (0)	307 (36)
PH TIC BN			
	SBP (mmHg), median (IQR)	91 (11)	128 (111, 142)
	HR (bpm), median (IQR)	297 (35)	87 (73, 101)
	GCS, median (IQR)	41 (5)	15 (12,15)
	Energy – low, n (%)	36 (4)	584 (68)
	MOI – blunt, n (%)	0	679 (79)
	Long bone fracture, n (%)	0	196 (23)
	Unstable pelvic fracture, n (%)	26 (3)	79 (9)
	Haemothorax, n (%)	0	116 (14)
	Abdominal bleeding, n (%)	26 (3)	81 (9)
	Lactate (mmol/L), median (IQR)	47 (6)	2.2 (1.3-3.5)
	PH crystalloid (ml), median (IQR)	11 (1)	0 (0-250)

[1]The COAST. score categorises this variable. The cohort values are reported as "model value, n (%)"
Bpm, beats per minute; SBP, Systolic blood pressure; GCS, Glasgow Coma Scale; HR, Heart Rate; CPR, Cardiopulmonary resuscitation; MOI, Mechanism of injury; #, Fracture; PH, Pre-hospital.

4.5.3 **Performance of the PH TIC BN**

4.5.3.1 Discrimination

In the complete patient sample (n=858) the 11 variable 'full PH TIC BN' exhibited reduced discrimination compared to the ED TIC BN (AUROCs: 0.92 (0.90-0.95) vs 0.94 (0.93-0.96), p = 0.03). In turn, the 10 variable 'no lactate PH TIC BN' had reduced discrimination compared to the full PH TIC BN (AUROCs: 0.90 (0.87-0.93) vs. 0.92 (0.90-0.95), p = 0.004) (**Figure 22**).

Overfitting, as assessed by inappropriate model optimism, was not demonstrated in the full PH TIC BN: AUROC external patient population (n=341) = 0.95 (0.91-0.99) vs internal validation cohort (n=517) = 0.90 (0.87-0.95), p=0.29.

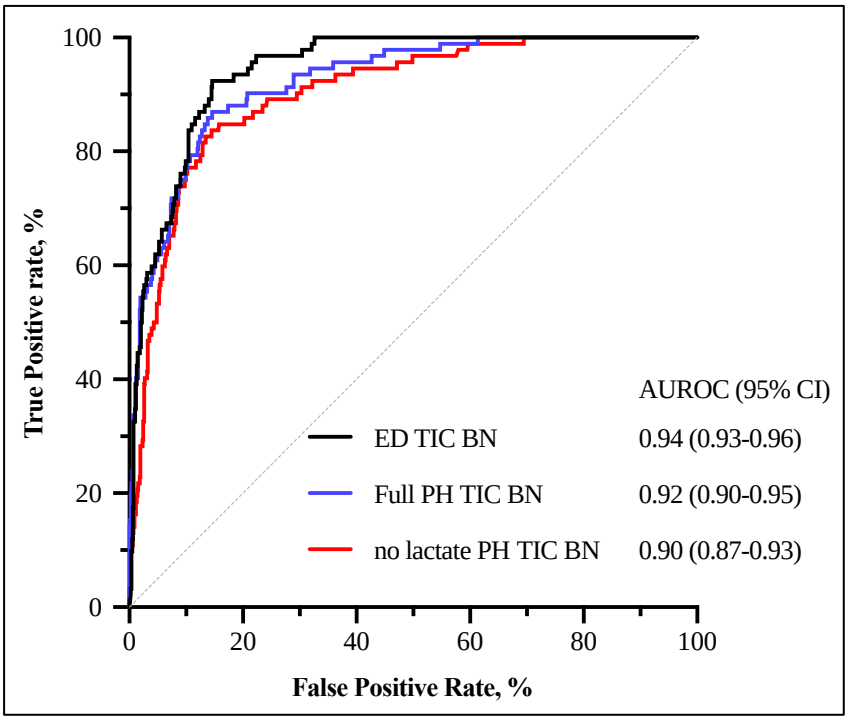

Figure 22: Discrimination performance of three variants of the Trauma Induced Coagulopathy Bayesian Network in 858 patients from the ACIT study. The receiver-operating characteristic (ROC) curves show the relationships between true-positive and false-positive TIC predictions in the original full ED BN (black line), the full PH BN (blue line), and no lactate PH TIC BN (red line). AUROC; Area under the receiver-operating characteristic curve, BN; Bayesian network, CI; Confidence interval, PH; Pre-hospital, ED; Emergency department.

4.5.3.2 Calibration

With both 10 and 11 input variables the PH TIC BN calibration plots have intercepts and slopes close to 0 and 1 respectively (**Figure 23**). In the no lactate PH TIC BN (**Figure 23B**) the model underestimates the probability of TIC at 5% to 65% risk and overestimates risk above 65%.

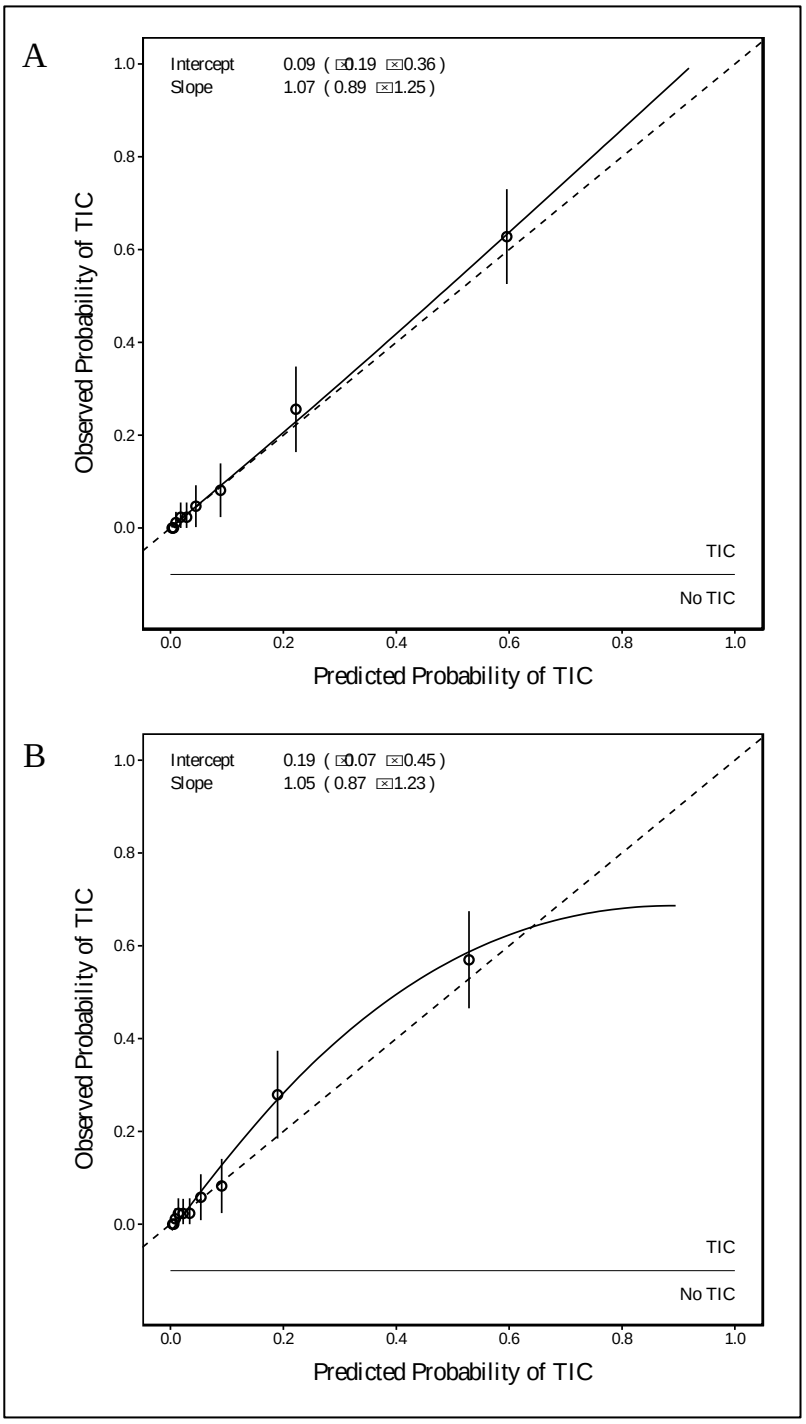

Figure 23: Calibration plots for the pre-hospital trauma induced coagulopathy Bayesian network.
The calibration plot demonstrates the relationship in 858 ACIT patients between perfect (dashed line) and observed (solid line) predicted values in
(A) full PH TIC BN
(B) no lactate PH TIC BN.

The rug plot at the bottom of each figure stratifies the distribution of predicted probabilities.

Circles with 95% confidence intervals represent deciles of patients grouped by predicted probability.

4.5.3.3 Overall performance measures – Brier Score

The PH TIC BN models had good overall accuracy. Degradation in overall model accuracy is demonstrated between the two versions of the PH TIC BN but not between the ED TIC BN and the full PH TIC BN. The Brier skill scores for the ED TIC BN, full PH TIC BN and no lactate PH TIC BN are 0.41 (0.32-0.50), 0.41 (0.32-0.50) and 0.31 (0.23-0.39) respectively.

4.5.3.4 Overall performance measures – net benefit

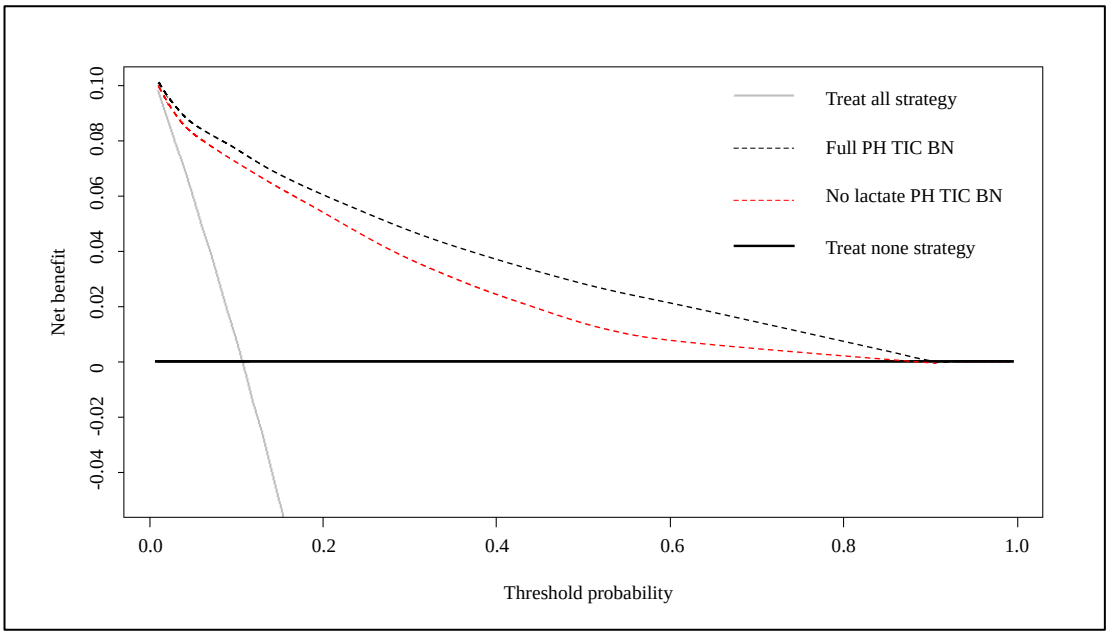

Figure 24: Decision curve showing net benefit for treating trauma induced coagulopathy in trauma patients. The net benefit of the full and no lactate PH TIC BN models are demonstrated against the default strategies of treat all and treat none. At any given threshold the model with the higher net benefit is the preferred model.

At all threshold probabilities the net benefit of using either PH TIC BN is greater than the default strategies. For example, if a threshold probability of 8% is used to designate an individual at high risk of TIC, the net benefit using of the full PH TIC BN over a strategy of treat all patients equates to five per 100, additional high-risk patients treated without increasing the number treated unnecessarily. When compared to a treat none strategy the model confers a benefit of eight per 100 additional patients treated.

4.5.4 **Performance of the PH TIC BN relative to other PH models**

Using complete case analysis for the COAST and PACT scores, the full PH TIC BN had significantly better discriminative ability than either the COAST (p=0.049) or PACT (p<0.001) scores (**Figure 25**). Removing the primary missing variable from the COAST and PACT scores and re-running the calculations resulted in poorer performance of both models (**Table 18**). The no lactate PH TIC BN was not significantly better than the full COAST score (p= 0.08) but was better than the full PACT score (p<0.001).

When model comparison was restricted to only include cases with full input variables for all three models only 293/858 (34%) patients remained. Under these conditions only 13/293 (4%) of patients had TIC. As a result, the true performance of the models became more uncertain: full TIC BN PH AUROC 0.85 (0.73-0.96), COAST 0.74 (0.58-0.89) and PACT 0.81 (0.73 to 0.90) and the models were not statistically different.

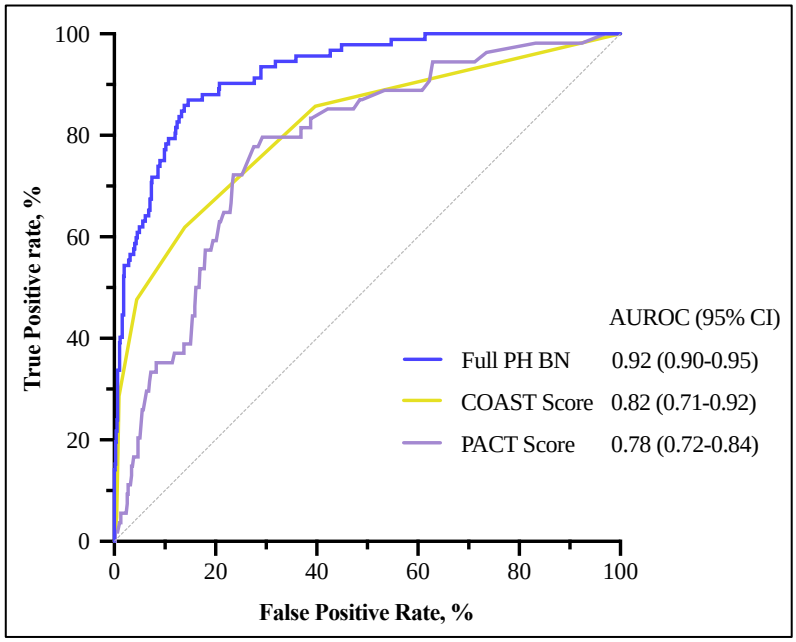

Figure 25: Discrimination performance pre-hospital TIC models in 858 ACIT patients. The ROC curves show the relationships between true-positive and false-positive TIC predictions in the full PH BN (blue line), the full COAST score (yellow line), and the full PACT score (magenta line). AUROC, area under the receiver-operating characteristic curve; CI, confidence interval.

Table 18: Discriminative ability of pre-hospital trauma induced coagulopathy models

	Full PH TIC BN	No lactate PH TIC BN	Full COAST score	No temp COAST score	Full PACT score	PACT score
n	858	858	452	743	526	814
AUROC	0.92	0.90	0.82	0.81	0.78	0.74
95% CI	0.90-0.95	0.87-0.93	0.71-0.92	0.75-0.86	0.72-0.84	0.69-0.80

PH; Pre-hospital, TIC; Trauma induced Coagulopathy, BN; Bayesian Network, SI; Shock index, AUROC, Area under the receiver-operating characteristic curve; CI, Confidence interval.

4.6 Discussion

4.6.1 Key Findings

The existing ED TIC BN can be adapted for use in the PH environment. The potential performance of this novel 11 variable PH TIC BN has been validated: it has an AUROC of 0.92 (0.90-0.95), is well calibrated, has good overall performance and has net benefit over two default treatment strategies. The predictive performance of the PH TIC BN was significantly better than other published models.

4.6.2 Interpretation of results

Removal of input variables from the ED TIC BN leads to degradation in the resultant model's performance. However, if the pragmatic assumptions made to calculate the performance of the PH TIC BN in his study are correct, the model appears to be promising and warrants further evaluation.

There are important caveats that arise from the assumptions made to calculate performance. As highlighted previously, the blood gas variables (BGV) (base deficit, lactate and pH) have the greatest impact on the ED TIC BN's result.[338] However, due to the features of a BN, Perkins demonstrated removal of all three BGV had minimal

impact on the model's performance. This finding is true when all other variables in the model are unchanged (BN with BGV AUROC: 0.95 (0.93–0.98) vs BN without BGV: AUROC 0.94 (0.91–0.98), p = 0.286). In this present study, not only are the BGV removed, but other variables are modified. The combination of removing key variables and modifying other variables has degraded the performance.

The full PH TIC BN has a better performance than the version of the model without lactate. As observed, only 15% of PH trauma patients at one of the two studied Air Ambulance sites had POC lactate measurement and only 4/22 Air Ambulance sites have access to POC lactate. In patients that do undergo PH lactate measurement, it is not clear whether a lactate obtained sooner after injury conveys the same amount of information than a lactate result obtained later. There is little available evidence to accept or refute this premise.[114] In this study, the ED lactate was substituted for the missing PH lactate. The impact of this substitution on the model's performance is not apparent.

4.6.2.1 The value of decision analytic results

With traditional performance metrics, it is possible to know a model's discrimination, calibration and overall accuracy results but not whether these values were sufficient to justify clinical use. For example, suppose there is a major bleeding model which predicts risk from 40-90%. This model has good discrimination and calculates higher predicted risk for bleeding patients. However, the model does not produce risk low enough to change clinical practice as clinicians decide to treat all patients with a risk greater than 40%. This threshold is a manifestation of the clinicians' judgement of the relative trade-off between errors of commission and omission. Variation amongst clinicians, patients and situations will influence which decision threshold is the most appropriate in a given situation.

DCA has some notable limitations. First, DCA assumes all treated patients will accrue the same benefit if they have the disease and suffer the same harm if treated without the disease. Second, DCA assumes that the threshold probability accurately summarises the costs and benefits of the intervention.[402]

4.6.3 **Strengths and limitations**

This study has assessed the feasibility and validated a BN model for PH TIC prediction. In a patient cohort of over 850 patients, the findings suggest that PH prediction of TIC is possible. The study is strengthened by adherence to the TRIPOD statement. The methods used to validate the model and report the study findings are consequently robust.

For example, model performance has been assessed in a group of patients external to the original development cohort. This step addresses the risk of overfitting. Overfitting of models results in exaggerated optimism of performance.[216] Overfitting occurs when a model function too closely fits a limited set of data points. In such cases, any noise in the data is misinterpreted as a signal.[407] Overfitting is best appreciated when the model is judged using an external dataset to the model development set.

Additionally, by comparing the PH TIC BN's performance against other existing prediction models in the same data set, the performance of the new model is contextualised. This useful comparison step is rarely undertaken in model validation studies.[408]

4.6.3.1 Limitations

The performance of the PH TIC BN was calculated using the same parameters as the ED TIC BN model. This study did not re-learn (or re-calibrate) the parameters of the model. Ideally the node probability tables, that describe the conditional probability relationships between variables in the model, should be re-learnt from data. Re-learning these parameters would account for the differences in the variables meaning between the ED and PH TIC BN models. For example, some variables in the PH TIC BN were measured at an earlier time point (Lactate, SBP, HR, Temperature, and GCS) but the probability distributions used for these variables were learnt from in-hospital data. In the case of the lactate variable, re-learning requires a dataset with PH lactate which was not available in this study. Re-learning the parameters may have improved the predictive performance of the model.

Substitutions in this study will also have impacted on the COAST and PACT models' performance. Four of the five COAST input variables required substitution as well as one of the six PACT variables. In this present external validation, COAST deteriorated from its previous small (n=133) external validation AUROC = 0.94 (0.88 - 0.99)[260] to 0.82 (0.71-0.92) and PACT from its prior external validation AUROC = 0.80 (0.72 - 0.88) [261] to 0.78 (0.72-0.84). Additionally, testing all three models against the outcome definition of just one of the models, will likely have degraded the performance of the other two models.

The impact of the missing data also affected the external validation and comparison of the three PH TIC models. For example, DeLong's method of correlated ROC curve comparison requires paired model values for each patient. For example, if missing data resulted in a patient, n, having a model prediction for the PH TIC BN model but not the COAST score, the models' performance cannot be compared for that patient. Missing data in the full COAST model reduced the number of patients available for paired analysis from 858 to 452. Fewer patients led to a wider confidence interval, which in turn affected the statistical comparison. In this manner, missing data in the COAST model affected the calculated performance of the 'no lactate PH TIC BN' model during their comparison; the 'no lactate PH TIC BN' model's discrimination changed from an AUROC = 0.90 (0.87-0.93) to AUROC = 0.88 (0.80-0.96). This impacted the no significant difference result between these two models (p= 0.08) and severely limited final comparison of the models using full case analysis for all models.

Another important limitation of this study relates to the way in which the pre-hospital injuries were coded. In this study an assumption was made that injuries above a certain severity threshold would be apparent to the clinician. However, the validity of this assumption is not known. As noted by the COAST authors, patients with occult injuries, are at risk of being underscored by any PH model reliant on injury information. The ability of the CDST user (in this case the pre-hospital clinicians) to accurately diagnose injuries pre-hospital requires further investigation.

4.6.4 **What remains unknown**

This study design was selected to assess the feasibility and scope the likely performance of a PH TIC BN. This was a necessary step to understand the predicted model performance before undertaking further research. However, this study does not advance the understanding of the real world impact of the model. It is unclear if the PH TIC BN will still perform more accurately than other models when used prospectively. It is unknown whether the PH TIC BN will influence clinical decisions and ultimately how changes in clinical decisions influence patient and health system outcomes. To gain insights into the potential of the model to augment decision making and thus understand the model's impact, prospective clinical evaluation is required.

4.7 Conclusion

This chapter aimed to identify the set of variables available for pre-hospital TIC prediction and 11 variables have been described. The performance of the PH TIC BN demonstrates that an individual's risk of TIC can be accurately predicted from clinical information in the pre-hospital environment. The PH TIC BN performs better than any published PH TIC prediction model and therefore warrants further prospective evaluation.

The next chapter assesses the real-world pre-hospital performance of the PH TIC BN. The chapter seeks to understand how the model will impact a clinician's risk assessment.

CHAPTER 5

THE EMERGENCY PRE-HOSPITAL ARTIFICIAL INTELLIGENCE IN TRANSFUSION AND TRAUMA INDUCED COAGULOPATHY (EMPHATTIC) STUDY

5.1 Scope of the chapter

This chapter describes the prospective multicentre 'Emergency pre-hospital artificial intelligence in transfusion and trauma induced coagulopathy' (EmPHATTIC) study. The EmPHATTIC study is a Phase I Impact Study of the PH TIC BN on pre-hospital decision making. The study focuses on the ability of the PH TIC BN to influence two pre-hospital decisions: is the patient at risk for trauma induced coagulopathy (TIC)? And will the patient receive a blood transfusion within four hours of injury?

5.2 Introduction

A pre-hospital healthcare professional has the opportunity to start therapy and potentially improve patient outcomes early after injury. This early treatment relies on the ability to identify patients most likely to benefit from the intervention. (see **2.5.2** and **3.5.1**) However, errors in decision making may lead to harm from either unnecessary treatment or delays in care. Errors such as these have profound consequences, yet important decisions are frequently based on incomplete information and uncertain risks. (see **2.5.1**, **2.5.2** and **2.5.3**)

TIC is the clinical manifestation of an endogenous multi-factorial and multi-mechanistic failure of the coagulation system, exacerbated by improper resuscitation techniques. (see **1.4**) The diagnosis of TIC is challenging pre-hospital. (see **2.5.2.1**) Coagulation assays in the pre-hospital environment (such as POC INR) have no proven benefit in the diagnosis or management of TIC.[122, 123] Viscoelastic haemostatic assays provide more information than CCTs, but are not yet widely available for pre-hospital use. (see **1.4.4.2**) Beside diagnostic assays, mathematical TIC prediction models have been used to risk-stratify patients. (see **1.6.3**) Unfortunately, neither TIC nor major haemorrhage prediction models are yet accurate enough to guide treatment decisions.[266]

Modern computational methods present an opportunity to develop powerful clinical decision support tools (CDST). CDSTs have the potential to augment clinical decision making and improve patient care.[310] Such models require careful development and validation in the target patient population. But model validation is not sufficient to provide evidence of a model's clinical impact. (see **4.5.3**) Without impact studies, the efficacy of a new CDST is unknown. Therefore, to introduce CDSTs safely into clinical practice requires impact studies.[271]

Very few studies describe the impact of CDSTs in clinical practice.[267] Compared to model development and validation, impact studies are difficult to conduct and expensive. These costs may be justified if there is evidence that the CDST might improve decision making in the target population. When performed well, impact studies

can provide real world data on CDST performance. Such studies can provide insights that guide the tool's implementation into clinical practice. (see **1.6.5**)

Currently no pre-hospital CDSTs for major haemorrhage or TIC have undergone impact analysis. Accordingly, most of the major haemorrhage and TIC CDST's remain research tools only. The real-world clinical impact of decision-support in this field is unknown.

5.2.1 A Bayesian Network model for pre-hospital TIC prediction

The previous chapter demonstrated that the PH TIC BN was able to predict a patient's risk of developing TIC using only pre-hospital clinical information. In a patient database evaluation of the PH TIC BN, the model had excellent predictive performance; discrimination AUROC = 0.92 (0.90 - 0.95), excellent calibration, good overall accuracy (BSS = 0.41 (0.32-0.50) and provided theoretical net benefit. Notably, the PH TIC BN performed better in the target population than the other published PH TIC prediction models (**CHAPTER 4**).

However, before the PH TIC BN model can be recommended for use in clinical practice, there are several unanswered questions. The conditions required for a model to have clinical use were outlined in **3.2.1.** Using this approach, the key remaining questions about whether the PH TIC BN is likely to be useful, have been summarised in **Table 19**.

Table 19: Remaining questions to assess the impact of the PH TIC BN and the required study design

Remaining questions	Answerable without affecting patient care
Does the PH TIC BN inform decision-making at a critical point in the treatment pathway?	Yes
What is the reliability of data capture for the model's inputs?	Yes
Is the model's output interpretable by the user?	Yes
Will clinicians use the PH TIC BN given a choice?	Yes*
Does the PH TIC BN enable the user to perform an accurate analysis of individual patients' risk?	Yes
Does the PH TIC BN maintain predictive accuracy when used prospectively and outside the limitations of the training dataset?	Yes
Can the model accurately predict the risk of TIC and need for PH blood transfusion?	Yes
What is the difference in predictive performance between the model and native clinical decisions?	Yes
Can the PH TIC BN alter clinicians' situational awareness, risk perception and decision making?	Yes
Does the PH TIC BN positively impact care?	No

* while aspects of this question can be assessed without affecting patient care, evaluating the uptake of the model in clinical practice may be best achieved outside of the research environment

The purpose of the present study is to address the key outstanding questions safely. To answer these questions requires a prospective study design. However, prospective use of the PH TIC BN may unexpectedly alter its predictive performance. By blinding clinicians to the model's output during patient care, many of the unanswered questions can still be addressed while minimising inadvertent patient harm. There are obvious drawbacks to a study design that blinds clinicians to the model's output during patient care. Such a study cannot address fundamental questions of whether the model improves the health system or patient outcomes or whether clinicians will use the model given a choice. To answer these questions will require future endeavours. Nevertheless, a stepwise approach to impact analysis aligns with the phased process suggested by the International Diagnosis and Prognosis Prediction group (IDAPP).[409]

5.2.2 **Phases of impact analysis studies**

Phase I, of the IDAPP's four-phase approach, is the 'exploratory phase'. The exploratory phase aims to determine whether the CDST is ready for full impact analysis. Assessment includes checking the CDST's inputs and output variables are sensible and comprehensible and re-checking the model's predictive abilities. By the end of the exploratory phase, the CDST should be finalised. Phase II, 'preparation phase' aims to define the CDST's delivery mode, study design, assess the acceptability of the CDST and feasibility of the impact study. Phase III, 'experimental phase' is when the effectiveness of the CDST on clinically relevant outcomes is undertaken. The final phase, 'long-term implementation phase' evaluates the translation of the CDST from a research setting into regular clinical practice.

5.3 Aims and Hypothesis

5.3.1 **Aims**

5.3.1.1 *Aim 1:* Prospectively assess the impact of the PH TIC BN model on pre-hospital clinicians' predictive performance to identify patients at risk of TIC and patients that will receive blood transfusion early after injury.

5.3.1.2 *Aim 2:* Compare the predictive performance of the PH TIC BN and expert clinical judgment to determine the risk of TIC and receipt of blood transfusion early after injury

5.3.1.3 *Aim 3:* Assess the reliability of data capture for the model's inputs

5.3.1.4 *Aim 4:* Explore the influence of the PH TIC BN on the clinician-reported difficulty in decision making

5.3.2 **Hypotheses**

5.3.2.1 *Hypothesis 1:* There is no difference in clinicians' discriminative performance before vs after receiving information from the PH TIC BN in identifying 1) patients at risk of TIC and 2) patients receiving a blood transfusion

5.3.2.2 *Hypothesis 2:* There is no difference between clinicians and the PH TIC BN in identifying 1) patients' risk of TIC and 2) patients receiving a blood transfusion

5.3.2.3 *Hypothesis 3:* All 11 PH TIC BN variables can be reliably captured by the time of the pre-hospital primary survey

5.3.2.4 *Hypothesis 4:* There is no difference in clinicians' degree of self-reported decision difficulty before vs after receiving information from the PH TIC BN

5.4 Methods

5.4.1 **Study Design**

The EmPHATTIC study was a prospective real-world phase I impact study conducted at two Air Ambulance sites. The study assessed the ability of the PH TIC BN to alter clinical judgement. Clinicians were asked for their prediction of a patient's risk of TIC or receiving a blood transfusion. Clinicians were also asked to report how difficult they found these predictions. To elicit clinician's predictions, interviews were conducted after each patient treatment episode on return to the headquarters using a standardised interview template (**Appendix B**: EmPHATTIC interview questions).

Clinicians' risk predictions were recorded at three different times: first, after the primary survey (Time Point 1); second, after the patient handover in the ED (Time Point 2); and finally, following the additional information of the PH TIC BN (Time Point 3) (**Figure 28**). All pre-hospital clinicians that attended the patient were eligible for interview. When more than one clinician's response was available, the mean value of the clinicians' responses was applied.

The PH TIC BN was used to compute predictions of the risk of TIC and the receipt of blood transfusion. Input variables for the PH TIC BN were entered from the information available immediately after the primary survey. Clinicians were given the model's prediction once they were no longer responsible for patient decisions (Time Point 3). Comparisons were made between the predictive accuracy of clinicians with and without the PH TIC BN information and between native clinical judgement and the PH TIC BN. Decisions were stratified by subjectively reported decision difficulty. The study enrolled clinicians for six months from the 07/01/2019.

5.4.2 The Pre-hospital trauma induced coagulopathy Bayesian network

Development and validation of the PH TIC BN are described in **CHAPTER 4**. The following sections describe how the model was used as a CDST in the EmPHATTIC study.

5.4.2.1 Predictor variables in the PH TIC BN model

In order for the PH TIC BN to guide decision making early after injury, the model needs to be used directly after the HEMS clinician's primary survey. To assess the feasibility of early model use the acquisition of each of the 11 input variables was analysed.

The PH TIC BN has up to 11 input variables (**Table 15)**. Input variables for the PH TIC BN were collected by the pre-hospital care clinicians and entered into the model by the Primary Investigator (PI) using the model's web interface (www.traumamodels.com). The PI and the clinicians were not aware of the patients' coagulation status at the time of PH TIC BN use.

Pre-hospital lactate measurement differed between the two Air Ambulance sites. LAA did not perform pre-hospital lactate measurement. Whereas, AAKSS advised its clinicians to measure the lactate in patients with a high clinical suspicion of bleeding using the Xpress Lactate device (Nova Biomedical, Waltham, Massachusetts, USA)

5.4.2.2 Design of the clinical decision support tool's output

The PH TIC BN's output is a probability. The probability describes the patient's likelihood of having TIC when they arrive in the ED. The clinical implication of the raw probability value is unlikely to be evident to the end-user of the model. To assist with comprehensibility, a visual presentation of the model's output was produced. The figure includes five risk categories (**Figure 26**). The width of each risk category represents the proportion of patients within the risk category. The proportions were derived from the 858 patients used in the **CHAPTER 4** validation cohort.

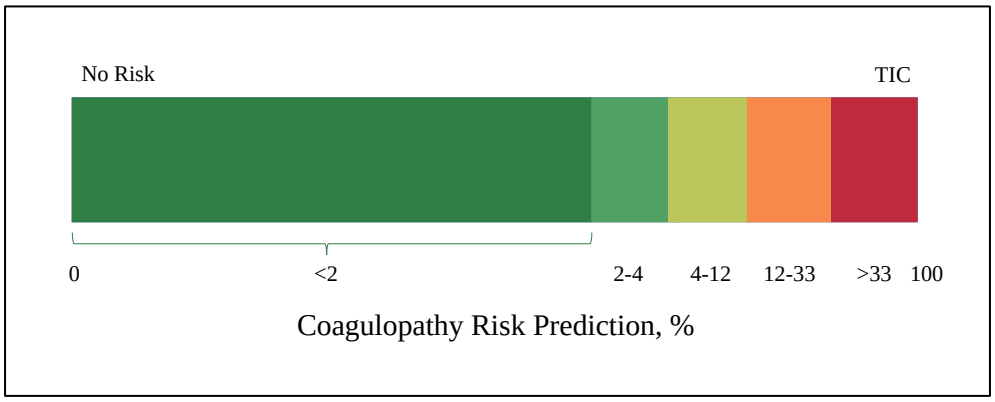

Figure 26: The PH TIC BN's output scale. The width of each of the five risk categories is proportional to the frequency of patients in the validation patient cohort with that risk prediction. The categories represent very low risk (dark green), low risk (light green), medium risk (yellow-green), high risk (orange), very high risk (red).

The output scale demonstrates 60% of the validation cohort had less than a 2% risk of TIC. The next four categories represent the risk of TIC in 60-69%, 70-79%, 80-89% and ≥90% of the validation cohort. The scale allows rapid visualisation of a patient's risk to help contextualise the clinical implications of the PH TIC BN result.

The categories are arbitrarily divided into five risk groups. The medium-risk group contains the risk thresholds for two methods to potentially dichotomise the PH TIC BN model's output. The first method uses Youden's index, which is maximised at a probability threshold of 10.8%. The alternative strategy if model dichotomisation was desired is to pick a probability threshold that produces the desired test sensitivity. For example, a 90% sensitivity occurs at a probability threshold of 6.6%. Both of these

potential operating point values (10.8% and 6.6%) lie in the 4-12% medium risk category. The scale in **Figure 26** rapidly communicates that patients with similar probabilities to these two thresholds place a patient in the top 70% risk of TIC of trauma patients.

5.4.2.3 Prediction of blood transfusion

To enable the PH TIC BN to predict blood transfusion, the model's TIC output was converted into blood transfusion volumes. Continuous blood volume prediction was unfeasible due to the relatively low number of patients receiving a blood transfusion in the validation population (**Table 12**). Therefore, clinically relevant categories of blood transfusion were sought.

The lack of clinical utility in the traditional definition of massive transfusion (≥10units of PRBCs in 24 hours) has been previously discussed (see **3.6.3.2**). To better reflect the early requirement for transfusions in bleeding patients, a shorter time interval was required. The interval chosen for the EmPHATTIC study was selected from an ongoing Delphi study seeking to gain consensus on a new definition of critical bleeding. This study suggests >4 units of PRBCs transfused within 4 hours of injury represents critical bleeding. A recent external validation of the Mina haemorrhage prediction model (**see 1.6.1.3**) used similar cut-offs.[231]

The amount of PRBCs patients received within 4 hours was assessed in the validation cohort. Validation cohort patients were split into three clinically relevant categories of PRBC [0, 1-4 or >4units] (**Table 12**). The distribution of PH TIC BN probability scores was assessed in each of the three transfusion categories for patients in the validation cohort (**Figure 27**).

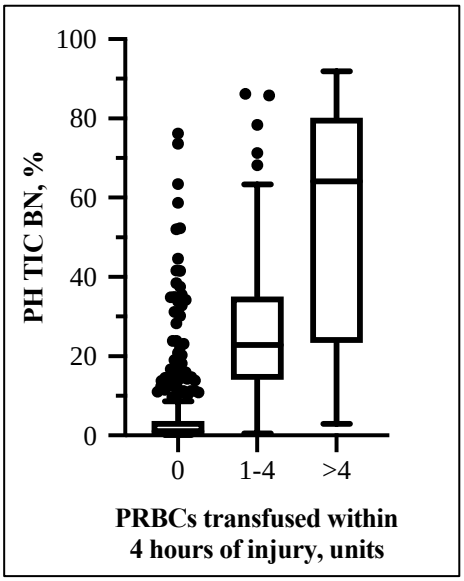

Figure 27: Distribution of PH TIC BN probability by the number of packed red blood cell units transfused within 4 hours of injury in the validation cohort patients. Tukey box and whiskers plot

A supervised learning algorithm was used to define the probability thresholds in the PH TIC BN output. The algorithm described which probability thresholds of the PH TIC BN were most closely associated with patients that received three categories of PRBC transfusion [0, 1-4 or >4units]. The algorithm optimised the relationship of the PH TIC BN output probability to the PRBCs unit categories using the Minimum Description Length Principle. This discretisation algorithm was run using the "Optimal Binning" function of SPSS 26 (IBM, Armonk, New York, USA).[410] Clinicians in the EmPHATTIC study were told the category of PRBC predicted by the PH TIC BN using the threshold values derived from this optimal binning algorithm (**Table 20**).

Table 20: Lower and upper PH TIC BN probability thresholds to define the number of transfused packed red blood cells within four hours of injury

PRBCs, units	PH TIC BN probability Lower threshold, %	PH TIC BN probability Upper threshold, %
0	0	15.32
1-4	15.33	52.85
>4	52.86	100

5.4.3 **Study population**

The study population consisted of pre-hospital consultants, registrars and paramedics from Air Ambulance Kent, Surrey and Sussex (AAKSS) and London's Air Ambulance (LAA).

5.4.3.1 Inclusion criteria

Pre-hospital doctors and paramedics who conveyed a patient to the hospital and provided written informed consent were included. Patients' were not enrolled into the study. Enrolled healthcare practitioners were asked for their predictions about adult (≥16 years old) trauma patients taken by AAKSS to either Kings College Hospital, Brighton and Sussex University Hospital, Southampton or St George's Hospital, or taken by LAA to the Royal London Hospital.

5.4.3.2 Exclusion criteria

Clinical encounters in which the pre-hospital clinicians did not remain with the patient until hospital admission were excluded. Clinicians were not asked to provide predictions for pregnant patients, those under 16 years old, patients with known bleeding diatheses including anticoagulant therapy (excluding 75mg Aspirin once a day) and patients with a greater than 12-hour interval between injury and clinical team arrival.

5.4.4 **Study Protocol**

During the study design in 2017, the Medicines and Healthcare products Regulatory Agency (MHRA) advised blinding clinicians to the model's results. The intention was to protect patients from potential harm. To blind clinicians and still assess model impact interviews were conducted following a patient treatment episode.

On returning to the air ambulance headquarters, clinicians were interviewed by the PI. Clinicians were asked for their prediction of the patient's probability of developing TIC by the time they arrived in the ED and receiving a blood transfusion within four hours

of injury. The interviews asked clinicians to recall their predictions of risk at two times: immediately following the primary survey (Time Point 1) and immediately following the patient handover in ED (Time Point 2). The patient's PH TIC BN prediction was calculated by the PI using the online model interface. The PH TIC BN's output was shown to the clinician using the linear scale (**Figure 26**) and category of blood prediction (**Table 20**). After this additional information clinicians were asked for their updated prediction of risk (Time Point 3) (**Figure 28**).

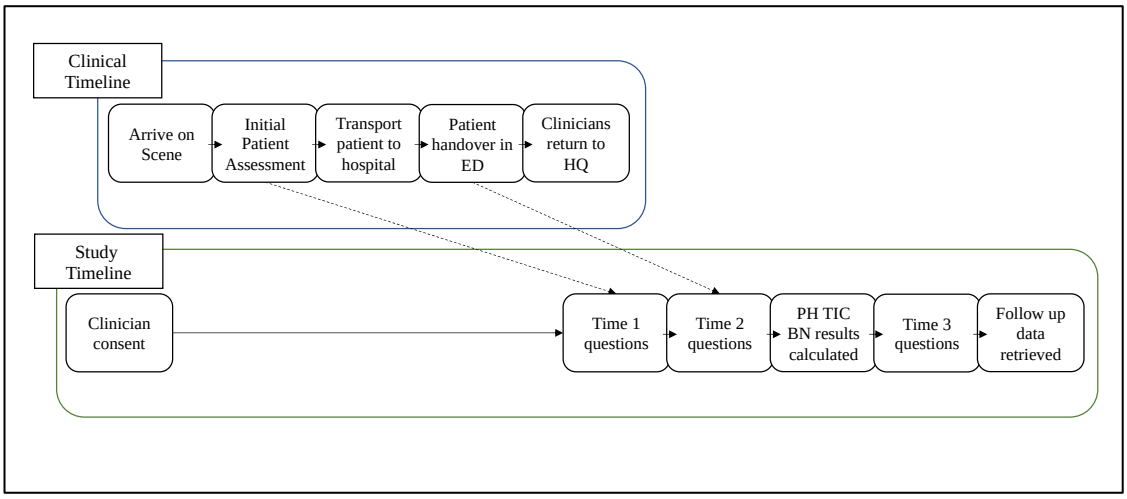

Figure 28: Clinical and study timelines Dashed arrows indicate the relationship between the clinical timeline and study time point questions.

Up to three clinicians were enrolled for each patient treatment episode. The typical clinical team consisted of a consultant, a registrar and a paramedic. To test the impact of the PH TIC BN in patients most likely to represent difficult decisions interviews were targeted to the most severely injured patients. High-risk patients were preferentially selected on the premise that a useful CDST is most likely to impact difficult decisions.

Clinician interview responses and anonymised patient data were collected and managed using the REDCap (Research Electronic Data Capture) (Vanderbilt, Nashville, USA) electronic data capture tool hosted at Queen Mary University of London.[411] Patient follow up data were abstracted from the receiving hospitals by the Air Ambulances as part of their standard clinical governance processes and added to the EmPHATTIC REDCap database. Hospital follow up data was retrieved and recorded to assess the

patient's true disease state and outcomes. Follow up data included a patient's hospital discharge injury diagnosis, laboratory coagulation results and the first 24 hours of blood transfusion requirements.

5.4.5 **Outcome classification**

The classification method used in the development and validation datasets of the TIC BN models was not possible in this clinical dataset. The previous classification utilised rich coagulation data gathered for the ACIT study. With this detailed data, an unsupervised clustering algorithm divided patients into normal or abnormal coagulation groups. The EmPHATTIC clinical dataset did not contain the same depth of coagulation assessment as the ACIT research study. As a result, TIC was classified as present when the admission prothrombin time ratio was greater than 1.2.[395] Where ROTEM values were available, thresholds of ExTEM CA5 ≤40mm, FibTEM CA5 <10mm and Maximum Lysis >15% were additionally used to classify TIC.[412] In cases of disagreement between these two laboratory test results, expert review of the patients was undertaken. Two TIC experts independently reviewed the clinical, laboratory, and thromboelastometry data of each discrepant case and applied the final coagulation status label. If no in-hospital coagulation parameters were available, the patient was excluded from further analysis.

5.4.6 **Sample size**

The study's sample size was intended to assess the impact of the model on clinicians' ability to identify patients at risk of TIC. Obuchowski's 2004 method[413-415] was used to calculate the number of patients required to detect a difference between clinicians with and without the PH TIC BN. Clinicians in this study were considered together as one predictor. Accordingly, the sample size calculation was only concerned with the number of patients required.

The sample size required is determined by a comparison of two paired ROC curves: clinical judgement with the PH TIC BN vs clinical judgement without the PH TIC BN.

The null hypothesis states there is no difference in performance. There is no empiric data available to estimate the native or supported predictive performance of clinicians in the pre-hospital diagnosis of TIC. The sample size calculation was therefore performed based on plausible estimated values. There is data to estimate the performance of the PH TIC BN on its own (**4.5.3.1**).

The estimated values are as follows: clinical predictive performance without the PH TIC BN was estimated to have an AUROC of 0.80; clinicians with the model would improve to an AUROC of 0.90; the correlation coefficient between the predictors was estimated at 0.6 for both patients with and without TIC; the ratio of normal coagulation to TIC patients was estimated at 4. Setting α = 5% and β = 20% the sample size required is 85 patients with the disease and 340 without the disease.

Given the time available for data collection and the number of severely injured patients treated by each air ambulance service, it was deemed unlikely that 425 patients would be attainable by one researcher. Despite the possibility of an underpowered study, valuable research outputs were anticipated from this exploratory study design regardless of the sample size.

5.4.7 **Statistical Methods**

The predictive performance of clinical judgement and the PH TIC BN were calculated using calibration, discrimination, and overall performance metrics when the outcome was binary. These methods were previously described in section **4.4.8**. The predictive performance of clinical judgement and the PH TIC BN on the categorical data relation to blood transfusion was assessed using overall classification accuracy and agreement analysis using Cohen's unweighted kappa (κ) statistic.[416] Overall classification accuracy was calculated as $\frac{\sum true\ positive + \sum true\ negative}{\sum total\ population}$. Cohen's κ assessed the level of agreement between the actual volume of blood received and the predicted volume.

The normality of continuous variables was assessed using Q-Q plots and the Shapiro-Wilk test. For calculations relating to subjective decision-making difficulty, Wilcoxon's

matched-pairs signed-rank test was used to compare continuous non-parametric data from the same observer and patient at different study time points. The Mann-Whitney U test was used to compare continuous independent non-parametric data. Numerical data are reported as median (IQR) and categorical data as frequency (n) and percentage (%). Statistical significance was set as a 2-tailed P value of <0.05.

5.4.7.1 Statistical computation and figure development

Statistical analysis and figure development were performed using SPSS 26 (IBM, Armonk, New York, USA), Prism 8 (GraphPad Software, San Diego, California, USA) and R statistical software (R Foundation for Statistical Computing; www.r-project.org version 3.6.0). The R packages were used "pROC" version 1.16.2[403], "ggplot 2" version 3.3.0[404], "rms" (Regression Modelling Strategies) version 5.1-4.[405] The BN's predictions were computed using the online interface at www.traumamodels.com.

5.4.8 **Ethics**

The EmPHATTIC study protocol and supporting documentation was submitted via IRAS 200742 for Research Ethics Committee (REC) approval. Given the nature of the study, the REC confirmed ethical approval was not required (REC reference 18/LO/2038). The study was sponsored by QMUL and gained prior approval from the Health Research Authority (HRA).

5.4.9 **Patient and Public Involvement**

The study protocol was presented at the Patient and Public Involvement meeting hosted by the South East Coast Ambulance Service on the 4th July 2018. The EmPHATTIC study outline was presented, and group discussions were facilitated. The attendees of the meeting were supportive of the study as it represented no harm to patients.

5.4.10 **Funding**

The study received no external funding.

5.5 Results

5.5.1 Baseline characteristics of clinicians and patients

A total of 51 clinicians were enrolled in the study, of which 21 were consultants, 14 registrars and 16 paramedics. The study included 194 complete patient-based interviews regarding 140 patients, of which 135 patients had in-hospital coagulation data, leaving 184 interviews for final analysis. The 184 interviews were conducted with a consultant in 56/184 (30%), a registrar in 75/184 (41%) or a paramedic in 53/184 (29%).

The patients had a median age of 31 years (IQR 23 - 47 years), median ISS of 17 (IQR 9 - 34) and 75% were male. Consistent with a selection towards the most injured patients, 26/135 (19%) had TIC and 62/135 (46%) received packed red blood cells within 4 hours of injury (**Table 21**). In patients with TIC, the relative risk of death was 55.0 (95% CI 7.5 - 402). The median PH TIC BN prediction was 6.6% (range 0.31% to 85%). There were a similar number of patients in each risk category (**Figure 29**).

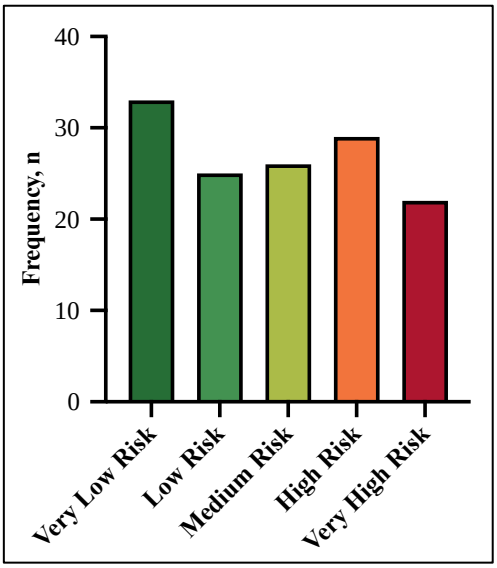

Figure 29: Histogram of PH TIC BN trauma induced coagulopathy prediction in risk categories communicated to clinicians

Table 21: Characteristics of patients included in the EmPHATTIC Study

	Missing data (n)	Total Population (n=135)	Normal Coagulation (n=109)	Coagulopathy (n=26)	P-value[a]
Patient Characteristics					
Age, years	1	31 (23, 47)	33 (23, 48)	27 (23, 45)	0.521
Male gender	0	101 (75)	87 (80)	14 (54)	**0.006**
Air Ambulance site 1	0	70 (52)	57 (52)	13 (50)	0.833
Injury to HEMS arrival, mins	0	32 (21, 44)	34 (22, 47)	29 (20, 35)	**0.034**
Injury to ED arrival, mins	1	88 (65, 114)	85 (66, 114)	101 (63, 115)	0.527
Injury severity score	3	17 (9, 34)	13 (9, 25)	38 (29, 45)	**<0.001**
Mortality	0	14 (10)	1 (1)	13 (50)	**<0.001**
Pre-hospital PRBC transfusion	1	38 (28)	20 (18)	18 (69)	**<0.001**
PRBCs within 4 hours, units	1	0 (0, 4.0)	0 (0, 1.8)	6.5 (3.0, 12.0)	**<0.001**
Pre-hospital Primary Survey					
Respiratory rate, bpm	9	20 (16, 25)	18 (16, 25)	22 (14, 28)	0.747
Heart rate, bpm	1	95 (80, 116)	91 (80, 110)	109 (56, 140)	0.143
Systolic blood pressure, mmHg	1	118 (91, 132)	120 (99, 139)	82 (52, 178)	**<0.001**
Glasgow coma scale	0	15 (9, 15)	15 (13, 15)	5 (3, 14)	**<0.001**
Blunt mechanism of injury	0	90 (67)	69 (63)	21 (81)	0.108
Low energy of injury	0	78 (58)	71 (65)	7 (27)	**<0.001**
Suspected haemothorax	0	34 (25)	26 (24)	8 (31)	0.465
Suspected abdominal bleeding	0	35 (26)	28 (26)	7 (27)	0.897
Suspected long bone fracture	0	29 (22)	21 (19)	8 (31)	0.199
Suspected unstable pelvic fracture	0	6 (4)	4 (4)	2 (8)	0.371
Any PH crystalloid	1	12 (9)	7 (6)	5 (19)	0.054
Pre-hospital lactate measured	0	15 (11)	12 (11)	3 (12)	>0.999
Pre-hospital lactate	120	2.3 (1.5, 3.7)	2.2 (1.5, 3.2)	3.7 (0.7, 3.8)	n/a
ED Arrival Blood Gas					
pH	27	7.3 (7.2, 7.4)	7.3 (7.2, 7.4)	7.1 (7.0, 7.2)	**<0.001**
Lactate, mmol/L	25	3.7 (2.1, 5.6)	2.8 (1.9, 4.7)	6.5 (4.0, 13.5)	**<0.001**
Base Deficit, mEq/L	32	4 (0.5, 8.8)	2.8 (0.2, 5.1)	9.5 (6.9, 16.5)	**<0.001**
Admission coagulation assays					
PTr	2	1.1 (1.0, 1.2)	1.1 (1.0, 1.1)	1.3 (1.3, 1.4)	**<0.001**
APTT (seconds)	25	24 (21, 28)	23 (21, 28)	31 (27, 41)	**<0.001**
Fibrinogen (g/L)	16	2.5 (2.1, 2.9)	2.5 (2.2, 2.9)	1.9 (1.4, 2.4)	**<0.001**
Platelet count ($x10^9$/L)	7	238 (190, 280)	251 (201, 296)	178 (111, 231)	**<0.001**
ExTEM CA5 (mm)	95	41 (32, 45)	42 (34, 48)	31 (26, 35)	**0.001**
FIBTEM CA5 (mm)	97	11 (7, 13)	12 (10, 14)	7 (4, 9)	**<0.001**
ExTEM Maximum Lysis (%)	102	6 (3, 8)	6 (4, 8)	3 (0, 30)	0.158

Data presented as number (%) or median (IQR). [a] P-value is the comparison between normal coagulation and coagulopathy groups. ED, Emergency Department; SBP, Systolic Blood Pressure; PRBC, Packed Red Blood Cells; CA5, Clot Amplitude at 5 minutes; MCF, Maximum Clot Firmness; INR, International Normalised Ratio; APTT, Activated Partial Thromboplastin Time;

5.5.1.1 Differences between air ambulance patient populations

Air Ambulance site 1 accounted for 21 of the interviewed clinicians, and 112/184 (61%) of the included interviews. Differences in patients' characteristics between the two study sites were present. Patients at site 1 were younger (27 years (20 – 40 years) vs 35 years (27 – 51 years), p = 0.002), more likely to have sustained a penetrating mechanism of injury (46% vs 21%, p = 0.005), had shorter times from injury to HEMS arrival (25 minutes (19 – 34 minutes) vs 37 minutes (30 – 57 minutes), p < 0.001) and shorter times from injury to ED arrival (77 minutes (56 – 104 minutes) vs 101 minutes (73 – 123 minutes), p = 0.002). Differences were present in methods of patient transport from the scene of injury to ED; site 1 transported 77% of their patients by road vs 39% from site 2, p <0.001.

There were no differences between the sites in the patients' ED arrival HR, SBP, temperature, pH or lactate or in the proportion of patients who underwent pre-hospital intubation and ventilation, major haemorrhage protocol activation, blood transfusion or administration of TXA.

5.5.1.2 Trauma induced coagulopathy classification

TIC classification was complicated by incomplete coagulation assay results. In 95 patients, there were no ROTEM data available. By far the most common cause for missing coagulation data was the assay not being run during the routine clinical care. In these 95 patients, their coagulation status was classified based on the admission prothrombin time ratio alone. In the remaining 40 patients, ROTEM data and prothrombin time ratio resulted in the same classification for 35 of the patients. The remaining five patients underwent expert review, of which 3/5 were classified as coagulopathic. One of the five cases for expert review required a third expert to reach agreement on the patient's classification designation.

5.5.2 Impact of the PH TIC BN on pre-hospital clinicians' predictions of the risk of TIC

5.5.2.1 Changes in clinician's risk predictions of TIC

The ability of clinical judgement to predict which patients would develop TIC on arrival to the ED varied between the time points. Immediately after the pre-hospital primary survey (time point 1), the performance of clinical judgement can be described with discrimination = 0.83 (0.74 – 0.92), calibration intercept = -1.19 (-1.73 - -0.65), calibration slope = 0.97 (0.50 - 1.44), and overall accuracy, given by the Brier skill score of 0.00 (-0.41 - 0.30). At time point 2, clinical judgment displayed a non-significant increase compared to TP1 in discrimination to 0.87 (0.80 – 0.95), p = 0.216. After clinicians were given the PH TIC BN results (time point 3), their discriminative performance remained largely unchanged; however, calibration and overall accuracy improved (**Table 22** and **Figure 30**).

Table 22: Predictive performance of clinical judgement in identifying patients at risk of trauma induced coagulopathy

	Time Point 1	Time Point 2	Time Point 3
Discrimination	0.83 (0.74, 0.92)	0.87 (0.80, 0.95)	0.88 (0.80, 0.95)
Calibration intercept	-1.19 (-1.73, -0.65)	-0.93 (-1.47, -0.40)	-0.79 (-1.38, -0.21)
Calibration slope	0.97 (0.50, 1.44)	1.16 (0.69, 1.63)	1.01 (0.58, 1.45)
Brier skill score	0.00 (-0.41, 0.30)	0.18 (-0.17, 0.44)	0.27 (-0.05, 0.51)

5.5.2.2 Performance of the PH TIC BN in predicting TIC

When used with data generated during the pre-hospital primary survey, the PH TIC BN's discriminative ability to identify patients at risk of TIC was 0.87 (0.79 – 0.95). The PH TIC BN's calibration intercept was 0.37 (-0.14 – 0.89), slope 1.11 (0.66 – 1.55), and overall model accuracy (BSS) was 0.34 (0.19 – 0.48) (**Figure 30**).

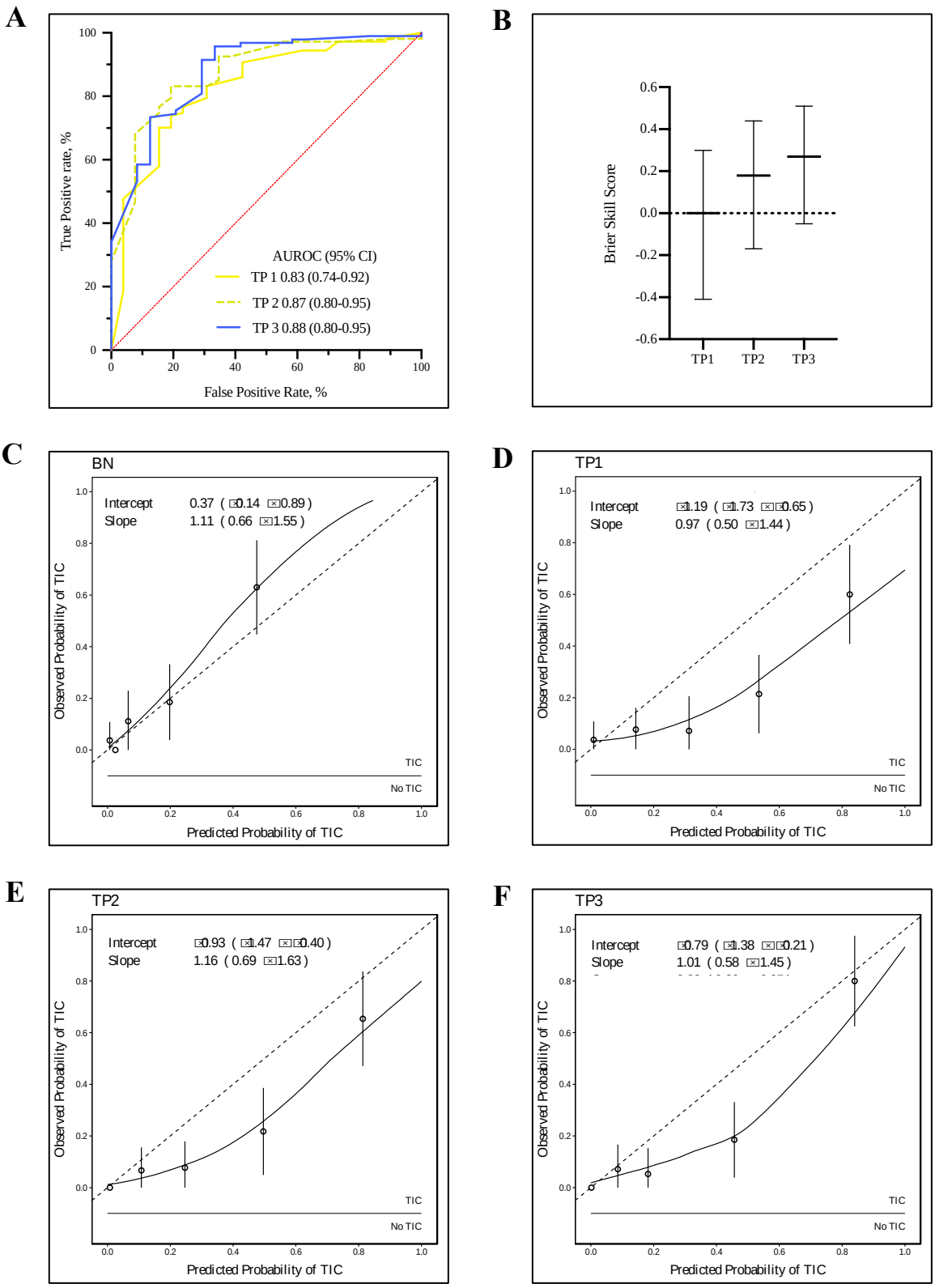

Figure 30: Predictive performance of clinical judgement and the PH TIC BN in identifying patients at risk for trauma induced coagulopathy.

Figure 30 (previous page): Predictive performance of clinical judgement and the PH TIC BN in identifying patients at risk for trauma induced coagulopathy.

A) Discriminative performance of clinical judgement at three time points illustrated with ROC curves. **B)** Brier skill score of clinical judgement at three time points. **C)** Calibration plot of the PH TIC BN and **D)**, **E)**, and **F)** Calibration plots of clinical judgement at study time points 1, 2 and 3.

Time point 1: immediately following the primary survey, solid yellow line; Time point 2: following patient handover off in the Emergency department, dashed yellow line; Time point 3: following PH TIC BN information, blue line. Brier skill score displayed with 95% confidence interval. The rug plots in figures C, D, E and F stratifies the distribution of predicted probabilities. Circles with 95% confidence intervals represent deciles of patients grouped by predicted probability.

AUROC, area under the receiver operating characteristic; BN, Bayesian network; TIC, trauma induced coagulopathy; TP, time point.

5.5.3 Impact of the PH TIC BN on pre-hospital clinicians' predictions of the need for blood transfusion

5.5.3.1 Changes in clinician's risk predictions of any blood transfusion

At time point 1, clinicians were able to identify patients that would receive any PRBC transfusion within four hours of injury with good discrimination 0.89 (0.84 – 0.95) and calibration. At time point 2, clinical discrimination had improved non-significantly compared to time point 1 to 0.92 (0.88 – 0.97), p=0.113. Following model information, the clinicians' ability to discriminate between patients that would receive PRBC transfusion was unchanged compared to ED handover (p>0.99) (**Table 23** and **Figure 31**).

Table 23: Predictive performance of clinical judgement in identifying patients at risk for blood transfusion within the first four hours of injury

	Time Point 1	Time Point 2	Time Point 3
Discrimination	0.89 (0.84, 0.95)	0.92 (0.88, 0.97)	0.93 (0.88, 0.98)
Calibration intercept	0.34 (-0.12, 0.80)	0.12 (-0.38, 0.63)	0.47 (-0.07, 1.01)
Calibration slope	1.21 (0.74, 1.69)	1.48 (0.91, 2.03)	1.63 (0.96, 2.30)
Brier skill score	0.47 (0.33, 0.60)	0.57 (0.44, 0.70)	0.58 (0.43, 0.72)

5.5.3.2 Performance of the PH TIC BN in predicting any blood transfusion

Immediately after the primary survey, the PH TIC BN's discriminative ability to predict any PRBC transfusion within four hours of injury was 0.89 (95% CI 0.83 – 0.95). The PH TIC BN's calibration intercept was 2.34 (1.91 – 2.77), slope 1.27 (0.87 – 1.67), and overall model accuracy given by the BSS was -0.04 (-0.26 – 0.15).

Figure 31 (next page): Predictive performance of clinical judgement and the PH TIC BN in identifying patients at risk for blood transfusion within the first four hours of injury.

A) Discriminative performance of clinical judgement at three time points illustrated with ROC curves. **B)** Brier skill score of clinical judgement at three time points. **C)** Calibration plot of the PH TIC BN and **D)**, **E)** and **F)** Calibration plots of clinical judgement at study time points 1, 2 and 3.

Time point 1: immediately following the primary survey, solid yellow line; Time point 2: following patient handover off in the Emergency department, dashed yellow line; Time point 3: following PH TIC BN information, blue line. Brier skill score displayed with 95% confidence interval. The rug plots in figures C, D, E and F stratifies the distribution of predicted probabilities. Circles with 95% confidence intervals represent deciles of patients grouped by predicted probability.

AUROC, area under the receiver operating characteristic; BN, Bayesian network; TIC, trauma induced coagulopathy; TP, time point.

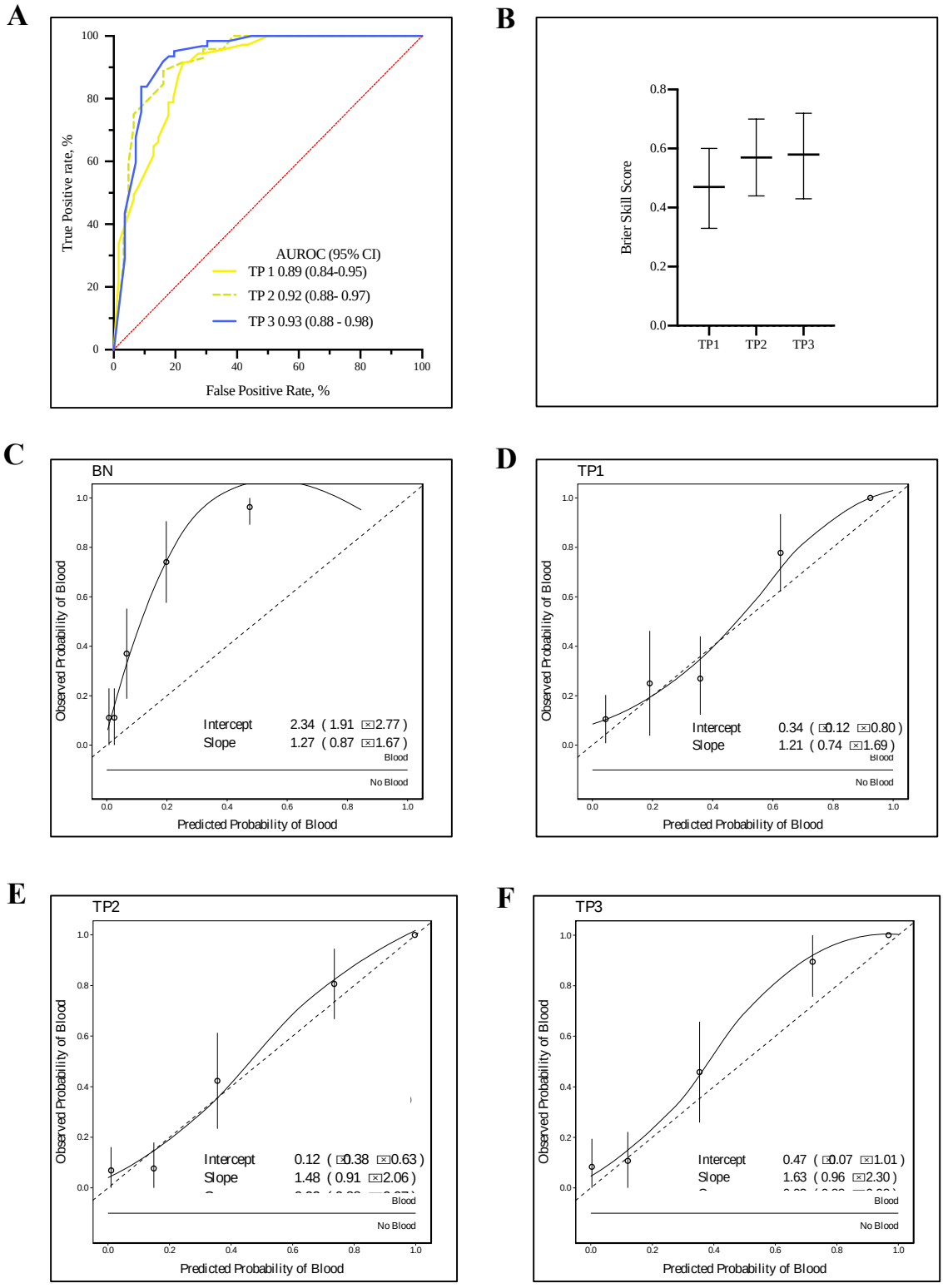

Figure 31: Predictive performance of clinical judgement and the PH TIC BN in identifying patients at risk for blood transfusion

5.5.3.3 Prediction of blood transfusion within clinically useful categories

Clinicians were asked to quantify the volume of blood they believed the patient would receive before (TP2) and after (TP3) they received additional PH TIC BN information. Two patients had incomplete categorical blood prediction data and were removed from further analysis. At TP2 clinicians had an accuracy of 66% and Cohen's κ = 0.473 (95% CI 0.356 - 0.589). At TP3 classification accuracy improved to 69% and Cohen's κ = 0.502 (95% CI 0.384 - 0.619) (**Table 24**).

Table 24: Confusion matrix of clinician classification of blood transfusion within 4 hours of injury before and after PH TIC BN information

		PRBC received, units		
		0	**1-4**	**>4**
PRBC predicted, units				
Clinicians without PH TIC BN (TP2)	**0**	**50**	3	0
	1-4	22	**23**	13
	>4	1	6	**15**
Clinicians with PH TIC BN (TP3)	**0**	**58**	5	0
	1-4	15	**19**	13
	>4	0	8	**15**

Results in green represent correct predictions, results in black are prediction errors. TP, Time point.

Visual analysis of the distribution of the PH TIC BN's probability output demonstrated reasonable discrimination between categories of blood transfused (**Figure 32A)**. However, the PH TIC BN under-estimated the volume of blood transfused (**Figure 31C** and **Figure 33A)**. Clinicians appeared to display better discrimination and calibration than the model (**Figure 31D** and **Figure 32B).**

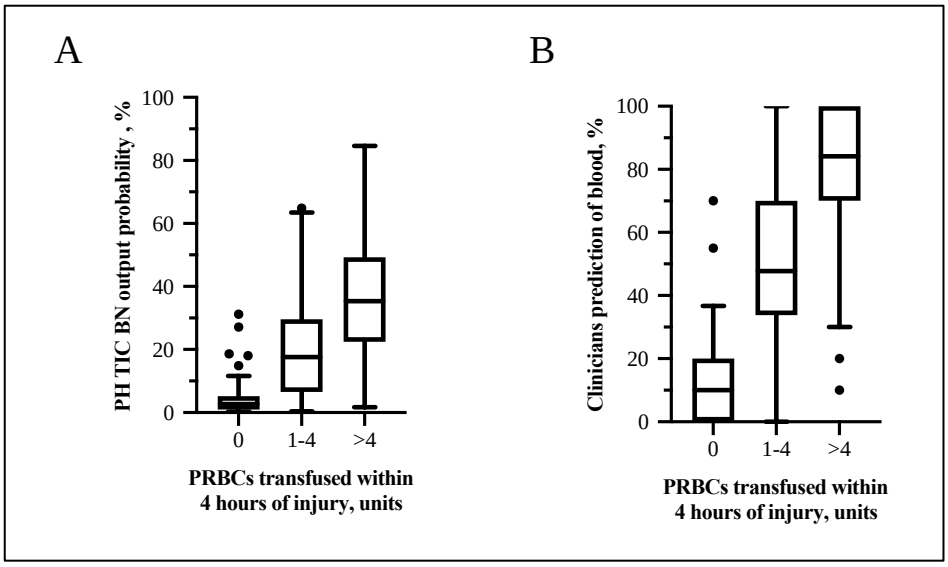

Figure 32: Distribution of blood transfusion predictions by categories of red blood cells transfused 4 hours after injury. A) PH TIC BN model B) clinical judgement. Tukey box and whiskers plot

Using the information available at TP1, the PH TIC BN predicted blood transfusion volumes with accuracy = 66% and Cohen's κ = 0.410 (95% CI 0.296 – 0.5) (**Table 25**).

Table 25: Confusion matrix demonstrating the predictive performance of the PH TIC BN to classify blood transfusion within 4 hours of injury

		PRBC received, units		
		0	**1-4**	**>4**
PRBC predicted, units				
	0	**69**	15	2
PH TIC BN[1]	**1-4**	4	**16**	22
	>4	0	1	**4**

Results in green represent correct predictions, results in black are prediction errors

[1] 1 response missing

5.5.4 **Comparison between clinical judgement and the PH TIC BN immediately after the primary survey**

5.5.4.1 Prediction of patients at risk of TIC

At TP1 there was no significant difference between the discriminative ability of the clinicians or the PH TIC BN to predict TIC (Clinicians: 0.83 (0.74 – 0.92) vs PH TIC BN: 0.87 (0.79 – 0.95), p = 0.330). Overall calibration given by the calibration intercept, was further from the ideal value of 0 in clinicians compared to the PH TIC BN (clinicians: -1.19 (-1.73 – -0.65) vs PH TIC BN: (0.37 (-0.14 – 0.89)). Similarly, clinicians' BSS value was further from perfect than the PH TIC BN (clinicians: 0.00 (-0.41 – 0.30) vs PH TIC BN: 0.34 (0.19 – 0.48)).

5.5.4.2 Predication of any blood transfusion

At TP1 there was also no statistical difference in the ability of the model to discriminate between patients that would receive any blood transfusion between clinicians and the PH TIC BN (clinicians: 0.89 (0.84 – 0.95) vs PH TIC BN: 0.89 (0.83 – 0.95) p = 0.874). However, clinicians appeared better calibrated. The PH TIC BN underestimated the probability of transfusion overall: calibration intercept clinicians: 0.34 (-0.12 - 0.80) vs PH TIC BN: 2.34 (1.91 – 2.77). Overall model accuracy was better for clinical judgement with a BSS 0.47 (0.33 – 0.60) vs PH TIC BN -0.04 (-0.26 – 0.15).

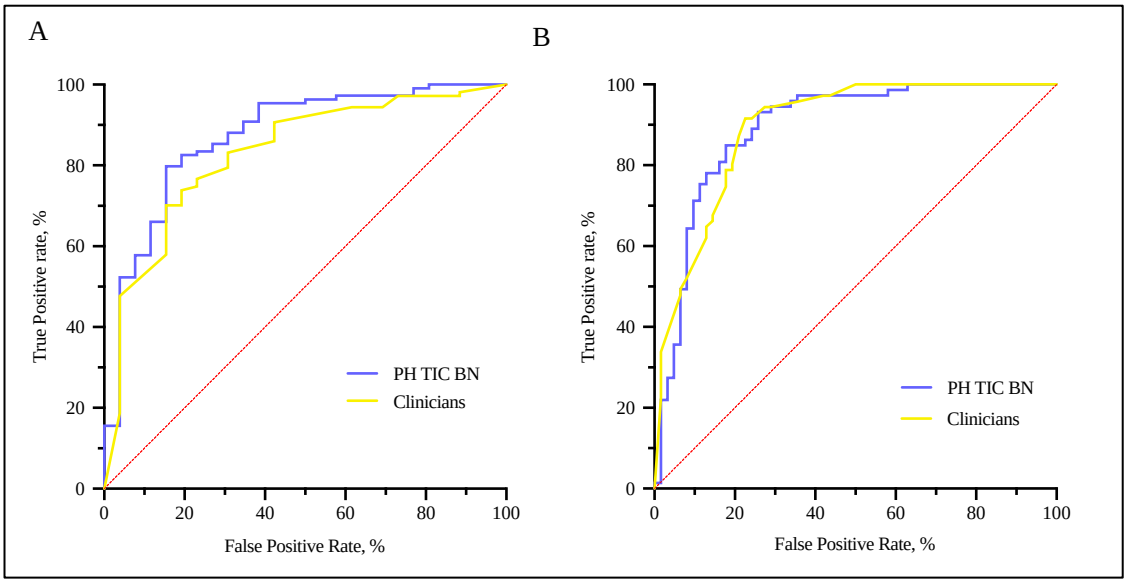

Figure 33: Comparison between discriminative ability between clinical judgement and the PH TIC BN. A) Trauma induced coagulopathy B) Blood prediction

5.5.5 Completeness of input data capture for the PH TIC BN

Immediately after the primary survey, all PH TIC BN input variables had 100% data acquisition except for energy of injury, SBP, and lactate. For one patient, the energy of the injury was unknown. In 4/135 (3%) patients, the SBP measurement was either unrecordable or unreliable. In the 15/135 (11%) of patients in which pre-hospital lactate was measured, only one measurement was available immediately post primary survey. In the other 14 patients, the measurement was taken before patient transport from the scene or en-route to the hospital. Lactate was measured 75 minutes (66 – 90 minutes) after injury, 18 minutes (12 – 34 minutes) after the primary survey and 36 minutes (23 – 48 minutes) before arrival in the hospital.

5.5.6 Influence of the model on decision difficulty

When clinicians reported how difficult it was to identify patients with TIC, a wide range of responses were collected. Overall, clinicians found the decision harder in patients with TIC (median difficulty 5.0 (IQR 1.3 – 7.0) vs without TIC 2.5 (IQR 1.0 – 6.0), p = 0.033). In paired analyses, the model statistically reduced decision difficulty both

overall (difficulty before model 2.5 (IQR 1.0 - 6.0) vs difficulty after the model 2.0 (IQR 0.5 - 5.0), p <0.0001) and in the cohorts of patients with and without TIC (**Figure 34**).

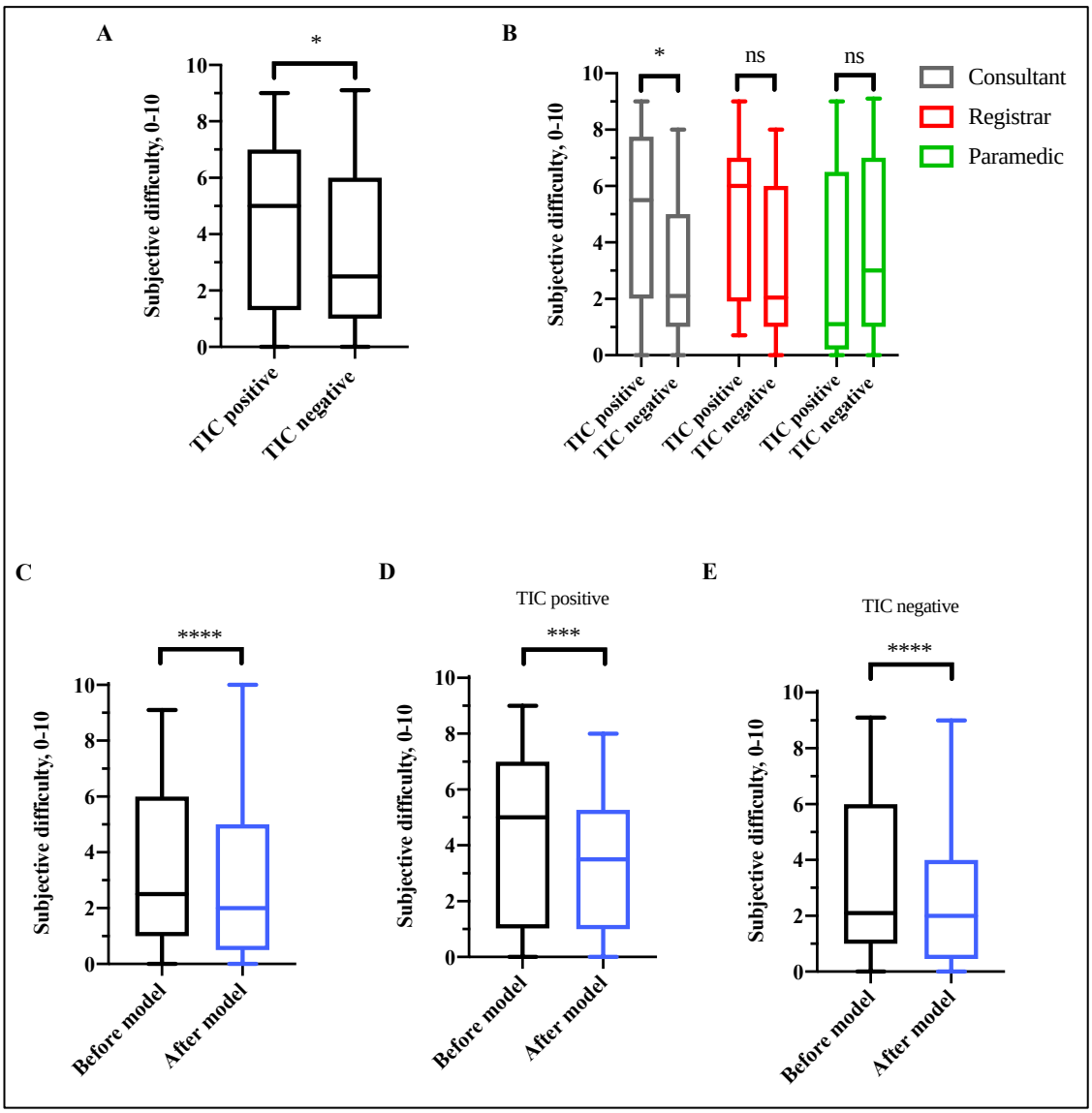

Figure 34: Clinician's self-reported difficulty scores in the prediction of TIC by clinician type. A) Difficulty stratified by the presence of TIC. **B)** Difficulty stratified by the presence of TIC and the grade of the clinician. **C)** Difficulty before and after the PH TIC BN model information. **D)** Difficulty in patients with TIC before and after the PH TIC BN model. **E)** Difficulty in patients without TIC before and after the PH TIC BN model.

* p<0.05, ***p<0.001, ****p<0.0001

5.6 Discussion

5.6.1 **Key Findings**

Clinical predictive performance did not significantly improve when supplemented with information from the PH TIC BN. However, there was evidence of a positive influence of the PH TIC BN on clinical judgement. The use of the PH TIC BN was associated with small improvements in the calibration and overall accuracy of clinical TIC prediction. However, the PH TIC BN's blood prediction output had suboptimal calibration with a systematic underprediction of blood transfusion.

Pre-hospital clinicians were better at identifying high risk patients at the end of the pre-hospital phase of care compared to the beginning. When performance was directly compared between the PH TIC BN and clinicians, it appeared that there was no difference in predictive performance immediately after the primary survey.

In this first prospective multicenter study of the PH TIC BN, all but one of the model's input variables were consistently generated during the primary survey. Pre-hospital lactate measurement was only available for 11% of the study's patients, and this missing data point led to a decrease in the PH TIC BN's predictive accuracy. The remaining key finding relates to the difficulty of decision making. Clinicians found decisions about patients with TIC harder, and despite little change in absolute values, decisions after decision support were less difficult.

5.6.2 **Impact of the PH TIC BN of clinical judgement**

This study's aim was to assess the impact of the model on clinical judgement. The null hypothesis relating to this study aim is accepted: there is no difference between clinicians' discriminative performance before or after augmentation with PH TIC BN information. The small differences in discriminative performance between TP2 and TP3 mean that failure to reach the required sample size has not affected this conclusion.

The failure to improve clinicians' predictive performance should be interpreted within the limitations of the study design. The impact of the model on clinical judgment was not assessed contemporaneously. The delay between generating a BN prediction and assessing augmented clinical judgment with the BN introduced a major confounder. Clinicians' predictive performance improved overtime without model augmentation. Effectively, this study assessed the impact of the model's prediction immediately after the primary survey on clinical judgement at the point of handover in the ED. This result may be more precisely interpreted thus: for expert pre-hospital clinicians, a prediction based on information from the start of the pre-hospital phase of care, does not augment clinical judgement when assessed at the end of pre-hospital care.

To improve future Bayesian Networks, the improvement in clinical performance over time is a useful finding. Ongoing patient observation may enable clinicians to identify influential cues from trends in a patient's physiology and their response to treatment. Such cues are not currently included in a single time point BN. These dynamic variables warrant investigation as additional inputs for inclusion in future models.

5.6.3 **Difference between clinicians and the PH TIC BN**

The study was not powered to detect a difference between clinical judgement and the model at time point 1, and thus, the finding of no difference may be underpowered. Nevertheless, the suggestion that the performance of the PH TIC BN may not be significantly different from clinical judgement at TP1 is an important finding. The clinical impact of a model is user and situation dependent. For example, the ability of a model to perform at the level of expert clinicians would be useful to those with less medical training. In austere environments such as mountain rescue or military contexts, such a model could support first responders. In these austere contexts increasing the prognostic performance of first responders could improve resource husbandry and patient triage within the whole trauma system.

It is also noteworthy, that there was no difference between the model and clinicians' performance at identifying patients that would receive blood transfusion. It is particularly remarkable because pre-hospital clinicians were responsible for pre-hospital

transfusion and consequently, they were responsible for part of the assessed outcome of blood transfusion within 4 hours of injury. Put a different way, pre-hospital clinicians' blood transfusion predictions were guaranteed to be correct in the 20% of patients that they gave a pre-hospital transfusion to. The PH TIC BN was not able to influence the outcome of blood transfusion in the same way and so its similar performance is a useful finding.

5.6.4 Data availability

The majority of the input data required for the PH TIC BN's computation is available early after injury. The lack of pre-hospital lactate in the majority of patients is the one exception. As the EmPHATTIC study only used data generated during routine care and clinicians previously discussed their uncertainty about the value of lactate measurement, (see **2.5.1.1**) it is apparent why much of the data was missing. It is also known from previous work that missing lactate data reduces the discriminative ability of the model (**Figure 22**). Future research is needed to investigate whether incorporating a pre-hospital lactate measurement into a model adds sufficient value to clinicians to justify the time measurement takes. Alternative approaches to determine a patient's perfusion status which are less logistically burdensome than measuring lactate would be desirable.

5.6.5 The role of the PH TIC BN in modifying the perception of difficult decision making

Decision support tools should address difficult decisions. The EmPHATTIC study found that patients with TIC were more challenging to predict than patients with normal coagulation. Somewhat surprisingly, the use of the model statistically reduced the degree of difficulty felt by clinicians in making decisions. The absolute reduction in difficulty after the model's prediction was small, and this is likely to be due to ongoing uncertainty. Clinicians were not informed of their patient's coagulopathy status after receiving the model's information. Thus, if the clinician disagreed with the model's prediction, they were left unsure whether to trust their initial prediction or the model's.

The interplay between decision difficulty and CDST design would be a valuable area to explore. A CDST that explains the impact of each variable on how the prediction is made might aid clinician's trust in the model. Increased understanding of the model's reasoning might reduce decision difficulty and free up clinician's cognitive "bandwidth". A study of the speed and quality of clinicians' decisions would be required to assess the impact of cognitive de-burdening.

5.6.6 **Strengths and limitations**

5.6.6.1 Strengths

The EmPHATTIC study was conducted with an exploratory intent. EmPHATTIC has provided insights into requirements for an interface between the model and the clinicians, evaluated the performance of the PH TIC BN and studied the design features in order to assess the potential impact of a CDST in the pre-hospital environment. These insights have been gained while maintaining patient safety. For example, the EmPHATTIC study has illustrated the approximate effect of the PH TIC BN on clinician's predictive performance. Such knowledge will be used to inform future sample size calculations for impact studies in phases II and III.

EmPHATTIC is the first study to prospectively assess the impact of a decision support tool of clinical judgement in pre-hospital trauma. The study demonstrated the input variables to the model were sensible, and the novel output scale was comprehensible to the user. The study aimed to include a sample of severely injured trauma patients to assess the PH TIC BN in its intended use population. This goal was met and is reflected by the inclusion of patients that received a pre-hospital blood transfusion (25%) and had TIC (20%). The study also enrolled over 50 clinicians from two air ambulance sites. By quantifying the performance of native clinical judgement, the study enabled the contextual determination of the PH TIC BN's performance. The multicentre study design also provided an assessment of the PH TIC BN in predominately both urban and rural pre-hospital systems and in patients with a mixture of mechanisms of injury, demographics and time from injury to assessment.

5.6.6.2 Limitations

Several factors limit the study. The most significant limitation results from the blinding of clinicians to the PH TIC BN result while treating a patient. This design led to an inability to assess the impact of the model on clinical decisions and the potential for recall bias. While the assessment of clinician's predictive performance is informative, changes in patient and health system outcomes depend on changes in clinicians' behaviour. Arguably the most important research question is whether the alternative *action (*prompted by the prediction of the CDST) *improve outcome* (final box **Figure 18**). This question requires a different study design with measurements at clinician, trauma system and patient levels of both processes and outcomes.

Another important group of limitations concerns the laboratory outcomes chosen to benchmark the PH TIC BN and clinicians. The previous method (see **4.4.6**) used to classify patients as coagulopathic was not possible in this clinical dataset. VHA is unusual in regular clinical practice at present; as a result, the majority of patients in this study are classified based on their prothrombin time ratio (PTr) alone. This leads to compounded errors as the PTr does not perfectly predict TIC. Future study designs will benefit from collecting complete coagulation data on each patient to ensure a comprehensive assessment of clotting status.

The second limitation connected with the PH TIC BN's output concerns prediction of transfusion requirement. There are conceptual issues as to whether the modelling goal is prediction of requirement or prediction of therapy. Prediction of the former i.e. the underlying physiological state of hypoperfusion (due to blood loss) is a superior aim in a modelling paradigm that uses causal relationships. However, transfusion requirement is linked to hypovolaemic state which is very challenging to objectively identify. Instead, a surrogate centred around blood transfusion has been used. This surrogate is associated with clinical error (over and under transfusion), data recording errors, and compounded by the logistic challenges in delivering the desired blood products to a bleeding patient.

As such, changes in clinical practice over time may be responsible for the underprediction of blood by the PH TIC BN seen in EmPHATTIC. The patients in the PH TIC BN's training dataset were treated between January 2008 and August 2013. The use of blood and blood products has undoubtedly changed during this period.[417] Blood transfusion now may be more liberal compared to a decade previously.

There are other contentious issues around the classification of clinically significant volumes of blood transfusion. This study considers three groups of PRBCs transfusion volumes within four hours of injury, but this is an arbitrary threshold. Tools such as the Critical Administration Threshold[391] and the Resuscitation Intensity[392] may account more appropriately for modern balanced transfusion practice.[390] However, the clinical use of these tools is minimal, and these newer concepts still rely on clinicians decisions to transfuse patients rather than identifying the need for therapy. Until hypovolaemia due to blood loss can be more accurately measured, a patient's need for blood transfusion will continue to be substituted for a prediction based on what patients in the training dataset received.

In the short term, adjusting the PH TIC BN's calibration for blood prediction provides a simple solution to improve blood prediction. Future BN modelling approaches may benefit by relating blood requirement as a child node to tissue perfusion, (**Figure 21**) rather than directly translating the prediction of TIC into the likelihood of blood transfusion.

The last notable limitation is the study's sample size. Due to limited time to collect patient data and the time involved in collecting follow up data from five MTCs, only 135 patients were analysed. This resulted in what would have been an underpowered estimate of the impact of the PH TIC BN on the clinical judgement of TIC. However, as the measured effect size was much smaller than anticipated, the sample size needed to find a difference would have needed to be much larger.

Additionally, and against the original intent, clinician interviews were rarely conducted with each of the three members of the clinical team. This precluded analysis of the differential effect of the model on each group of clinicians (paramedic, registrar and

consultant). This information was lost primarily due to the duration of each interview and the competing interests of the clinical team in preparing for their next mission. In EmPHATTIC, each patient generated 1.4 clinician responses on average. Some authorities believe that the aggregation of information in groups results in decisions that are better than of any one individual.[418] It is feasible that averaging clinicians' predictions improved the measured performance of "clinical judgement". Future studies may benefit from direct clinician data entry, rather than relying on interviews to elicit clinician's predictions. This would allow a detailed analysis of the impact of the model on clinicians with differing levels of experience and training.

5.6.7 **The EmPHATTIC results in context**

There is a decrease in the performance of the PH TIC BN when comparing the result of the internal-external validation in **CHAPTER 4** and the prospective external validation in this current study. In EmPHATTIC the PH TIC BN has reduced discrimination in the prediction of TIC (AUROC: 0.92 (0.90 – 0.95) vs 0.87 (0.79 – 0.95), p = 0.298), less precise calibration (intercept 0.09 (-0.19 – 0.36) vs 0.37 (-0.14 – 0.89)) and reduced overall accuracy (BSS: 0.41 (0.32 – 0.50) vs 0.34 (0.19 – 0.48). Although the discrimination comparison is not significant, the general trend across all the performance metrics suggest there is a degradation in performance. There are several plausible explanations for this result that require further investigation.

One possibility is the results of **CHAPTER 4** overestimate the model's performance. This could be due to overlap between the development and internal-external validation patient cohorts, with subsequent optimistic assessment of performance. Also, the assumptions relating to the reliability of pre-hospital diagnosis in **CHAPTER 4** may be incorrect. If clinicians are less accurate in their diagnosis than previously assumed, the model will appear to perform better. The reliability of pre-hospital clinical diagnosis requires further analysis. Alternatively, the seemingly reduced performance of the PH TIC BN in the EmPHATTIC study may simply be due to the small number of patients and wide confidence intervals.

The EmPHATTIC study joins a limited group of studies that aims to both assess a model's real-world impact on clinical judgement and improve reporting and transparency through adherence to reporting guidelines. A recent systematic review explored the application of AI to medical imaging; arguably the most developed area of medical AI. The systematic review compiled studies comparing expert clinicians to deep learning algorithms. The study found just two published randomised clinical trials and eight ongoing trials. Of 81 non-randomised clinical trials, only six were prospective real-world studies. The study found the median number of expert clinicians used as a comparison was 4 (IQR 2-9) and between the studies, there was less than 50% adherence to 12 of 29 TRIPOD items. The authors conclude that the risk of bias in the majority of published studies is high and the conclusions are often overstated.[419]

5.6.8 **What remains unknown**

The EmPHATTIC study was a phase I impact study in which clinicians were blinded to the PH TIC BN's prediction during pre-hospital treatment. As a result, important questions about the impact of the PH TIC BN on clinical decision making and patient and health system outcomes remain unknown. Furthermore, the EmPHATTIC study did not present a CDST to clinicians. Instead, inputs and outputs of the PH TIC BN were communicated to clinicians by a third party familiar with the model.

The clinical impact of a CDST depends on a clinician's decision to use the model and reception to its outputs in considering an alternative clinical decision. Future questions remain on how to incorporate the PH TIC BN into a CDST effectively. The goal is to seamlessly integrate the CDST into the clinical workflow, without adding burden to the clinicians. The model's output must also be presented to clinicians in such a way that the user can understand the clinical relevance. Ease of the CDST's use, the clinician's understanding of the output and her trust in the model's recommendation will all affect its impact and have not yet been assessed.

It is not clear whether the current version of the PH TIC BN can meaningfully augment decision making amongst expert pre-hospital clinicians. As these clinicians demonstrate improved predictive performance with time spent treating and managing their patient's

conditions. The model's predictive performance could be improved by changing the inputs. Variables which may hold substantial signal include metrics that encompass physiological trends and response to treatments such as blood transfusion. The model's performance may also improve with technology that can accurately measure perfusion status, rather than employing standard descriptors of shock status such as heart rate and blood pressure.

5.7 Conclusion

The EmPHATTIC study is the first prospective multicenter impact study of decision support on pre-hospital TIC prediction and blood transfusion in trauma. It is also the first study to quantify the predictive performance of clinical judgement in these areas. The PH TIC BN did not improve clinician's predictive performance at the end of the pre-hospital care phase but may be more influential immediately after the primary survey. The model may also be of use to non-expert first responders in austere environments.

Future studies are needed to address fundamental unanswered research questions. Phase II and III impact studies should assess whether this decision support improves outcomes. Such studies will require a CDST for clinicians to directly interact with the model and unblinding of clinicians to the model's prediction during patient care. Given the results of the EmPHATTIC study and re-interpretation of current MHRA regulations, a future trial design which randomises clinicians to CDST support or not during patient care is now a realistic possibility.

The next chapter explores the reliability of pre-hospital injury diagnosis to understand potential limitations to the performance of the model.

CHAPTER 6

THE RELIABILITY OF PRE-HOSPITAL INJURY DIAGNOSIS

6.1 Scope of the chapter

This chapter explores the reliability of pre-hospital injury diagnosis. A retrospective study is performed to compare clinicians' pre-hospital injury diagnosis with the patient's final injury diagnosis on hospital discharge. The chapter explores whether errors in pre-hospital injury diagnosis account for the decrease in performance of the PH TIC BN seen in the EmPHATTIC study.

6.2 Introduction

Prediction models that incorporate injury information are dependent on the accuracy of injury diagnosis for optimal model performance. In the EmPHATTIC study (**CHAPTER 5**), the performance of the PH TIC BN was inferior to the performance of the same model validated retrospectively (**CHAPTER 4**). The deterioration in model function may be due to a decrease in accuracy of the model's input variables.

In **CHAPTER 4**, the performance of the PH TIC BN was approximated by making assumptions about the accuracy of pre-hospital injury diagnosis. These assumptions were required as the patient database used to calculate the model's performance did not include pre-hospital injury diagnosis. **CHAPTER 4** assumed that injuries in the chest and abdomen were reliably diagnosed when the Abbreviated Injury Scale (AIS) was ≥3 (see **4.4.3**). The chapter also assumed unstable pelvic fractures would be accurately diagnosed. Little published data is available to support these assumptions. Overconfidence in the ability to accurately diagnose injuries pre-hospital may have led to inflated approximations of the PH TIC BN's performance.

6.3 Aims and Hypothesis

6.3.1 **Aims**

6.3.1.1 *Aim 1:* Assess the reliability of pre-hospital injury diagnosis for important predictor variables in the PH TIC BN model

6.3.1.2 *Aim 2:* Assess whether severe injuries are more accurately diagnosed

6.3.1.3 *Aim 3:* Identify factors which lead to deterioration in diagnostic accuracy

6.3.2 **Hypotheses**

6.3.2.1 *Hypothesis 1:* Pre-hospital injury diagnosis is sensitive and specific

6.3.2.2 *Hypothesis 2:* Severe injuries have increased diagnostic accuracy compared to less severe injuries

6.3.2.3 *Hypothesis 3:* Decreased consciousness is associated with decreases in diagnostic accuracy

6.4 Methods

A review of all patients conveyed to the Royal London Hospital by London's Air Ambulance (LAA) between 30/09/2017 and 22/04/2018 was undertaken. Two databases were compared to evaluate the accuracy of pre-hospital injury diagnosis. The first database contained the AIS injury scores from the pre-hospital diagnosis and the second database contained the AIS injury scores from the patient's hospital discharge diagnosis.

The pre-hospital injury diagnosis is coded for all patients assessed by LAA, using only information obtained during the pre-hospital phase of care. PH AIS scores are generated within 24 hours of a patient's assessment as part of a routine clinical governance process. The pre-hospital medical documentation (patient report form) is reviewed by a group of PH doctors, including at least one pre-hospital consultant. The AIS scores are entered into the database by a PH clinician reading the patient report form and scoring the injury. The PH clinician uses the AIS handbook to score the injuries and every diagnosed injury is recorded in the PH database. The AIS scores are agreed between the clinicians before finalisation.

The final diagnosis injury AIS scores were generated on patient discharge by trained trauma coding administrative staff. The final injury score was generated using the patient's electronic notes and included information derived from imaging and intra-operative findings. This injury data is stored in a locally maintained hospital database. Both databases coded the patient's injuries using the 2005 abbreviated injury score (AIS) system.[420]

Injuries were assessed in three pertinent body regions; chest, abdomen and pelvis. Injuries were classified as correctly diagnosed pre-hospital when within one AIS score of the final diagnosis. A tolerance of one point on the AIS scale was arbitrarily selected to reflect the clinical significance of injuries two AIS points apart.[256] Performance metrics were evaluated in the standard way. A subgroup analysis of all patients with a hospital discharge diagnosis of AIS ≥3 was performed to assess whether increasing injury severity was associated with improved diagnostic accuracy.

Patient records were matched between databases using an anonymised unique identifier in both databases. Only patients underwent the rapid review process and could be matched between databases were included. This work was conducted as a service evaluation and registered with the Barts Health NHS Trust Clinical Effectiveness Unit.

To investigate the effect of reducing Glasgow Coma Scale (GCS) on the risk of missed injuries, patients were grouped into three clinically relevant groups of GCS (<15, <8, 3). The baseline risk was calculated in patients with a GCS of 15, as the risk of a missed injury as a proportion of all patients with an injury. Risk in each GCS group were calculated relative to the baseline risk (i.e. patients with a GCS =15). Each GCS group was compared to the reference group with Fisher's exact test.

6.5 Results

During the study period, 688 trauma patients were attended by LAA and 177 met the study's eligibility criteria (**Figure 35**). Patients were typically young men, injured by a blunt mechanism of injury and had a median final ISS of 9 (**Table 26**).

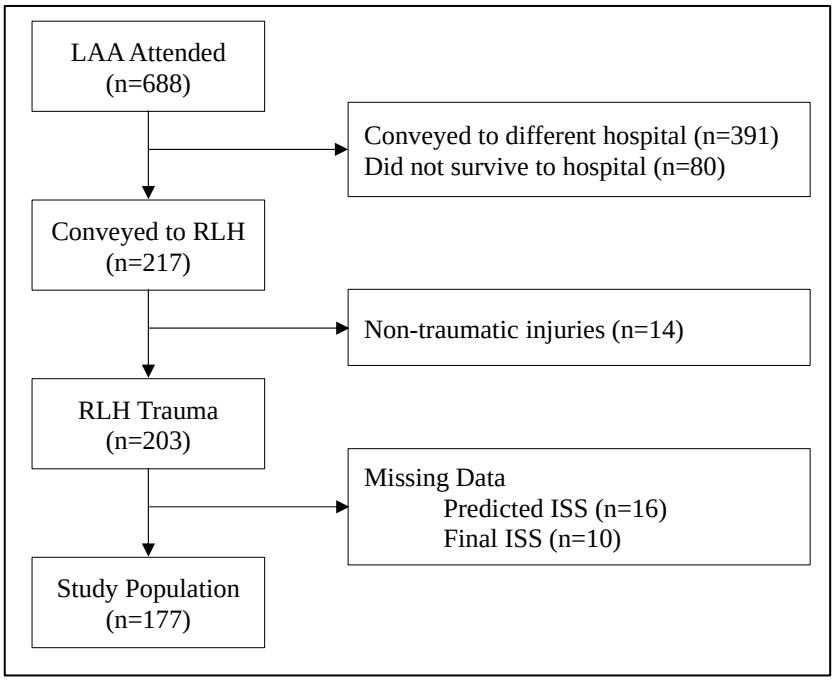

Figure 35: **Patient flow diagram.** LAA; London's Air Ambulance, RLH; Royal London Hospital, ISS; Injury severity score

Table 26: Characteristics of patients included in the study

	Missing data	Total population
n		177
Age, years, median (IQR)	1*	29 (20-42)
Male gender, n (%)	0	149 (84)
Blunt, n (%)	0	111 (63)
Pre-hospital ISS, median (IQR)	0	4 (2-10)
Hospital discharge ISS, median (IQR)	0	9 (1-17)
Suspected head injury pre-hospital n (%)	1	38 (21)
Initial pre-hospital GCS	46	15 (14-15)

GCS, Glasgow coma scale; ISS, Injury severity score; IQR, Interquartile range
*does not include 12 approximated ages

Overall, pre-hospital injury diagnosis had a sensitivity and specificity of 62% and 93% respectively. In total, 51/177 (29%) patients had one or more errors of pre-hospital injury diagnosis. Pre-hospital diagnostic error was more commonly due to an under-appreciation of injury severity, rather than over diagnosis. The pre-hospital diagnosis of chest injuries was the most sensitive (77%) and pelvic injury the least sensitive (35%). Specificity decreased from 96% in the pelvis to 90% in the chest. Subgroup analysis of patients with an AIS equal to or greater than three did not result in improved diagnostic accuracy (**Table 27**).

The proportion of missed injuries in patients with a normal GCS was 19/40. An inverse association was observed between the GCS and the risk of underscored pre-hospital injuries (**Figure 36**). However, none of the comparisons to the reference group (GCS=15) were significant at $p<0.05$ level. No association was found between the time of the shift (day or night) and the proportion of errors in pre-hospital diagnosis ($p=0.867$).

Table 27: Pre-hospital reliability of injury diagnosis

	Chest	Abdomen	Pelvis
All patients			
Patients with a final diagnosis AIS ≥1, n (%)	57 (32)	30 (17)	15 (8)
Performance of clinical judgement			
Patients with any error, n (%)	24 (14)	24 (14)	16 (9)
Final AIS - PH AIS, median (IQR)[1]	0 (-2, 2.75)	2 (-2, 2)	2 (-2, 3)
Sensitivity, % (95% CI)	77 (64, 88)	48 (29, 67)	35 (14, 62)
Specificity, % (95% CI)	90 (84, 95)	93 (88, 97)	96 (91, 98)
Positive Predictive Value (95% CI)	77 (66, 86)	58 (41, 74)	46 (25, 69)
Patients with an AIS ≥3			
Patients with a final diagnosis AIS ≥3, n (%)	31	10	10
Performance of clinical judgement			
Sensitivity, % (95% CI)	68 (49, 83)	40 (12, 74)	30 (7, 65)
Specificity, % (95% CI)	97 (92, 99)	99 (96, 100)	99 (96, 100)
Positive Predictive Value (95% CI)	81 (63, 91)	67 (29, 91)	60 (22, 89)
Analysis of diagnostic errors in patients with AIS ≥3			
Under-scored injury, n (%)	3 (10)	4 (40)	0
Missed Injury, n (%)	6 (19)	2 (20)	7 (70)
Over-scored injury, n (%)	4 (13)	1 (10)	0
Injury suspected but none present, n	0	1	2

[1] In patients with more than 1 AIS score between the final and pre-hospital diagnosis. When positive the (final AIS – PH AIS) value demonstrates an underdiagnosis of injury pre-hospital.
AIS, Abbreviated injury score; PH, Pre-hospital

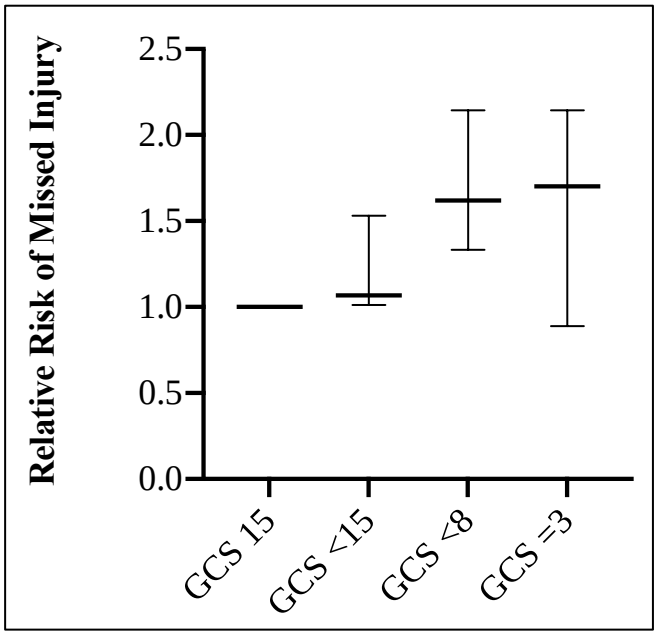

Figure 36: Relative risk of missed injury by the patient's initial pre-hospital Glasgow coma scale. Risk of missed injury is relative to diagnostic performance in patients with a GCS = 15. Missed injuries are all PH injuries diagnoses ≥2 AIS scores lower than the patient's injury diagnosis on hospital discharge. Thick bar, median relative risk value; whiskers, interquartile range of relative risk. GCS, Glasgow Coma Scale. GCS 15 group (reference group) includes 71 injuries, GCS <15, 36 injuries, GCS < 8, 15 injuries, and GCS = 3, 9 injuries. None of the reducing GCS groups are statistically different from the reference group.

6.6 Discussion

This study explores whether errors in pre-hospital injury diagnosis account for the decrease in performance of the PH TIC BN seen in the EmPHATTIC study. When averaged across the chest, abdomen and pelvis, the sensitivity of pre-hospital clinical diagnosis was approximately 50%. Surprisingly, in patients with more severe injuries, there was no improvement in diagnostic accuracy. Thus, previous assumptions made in **4.4.3** about increased diagnostic performance with more severe injuries led to an overestimation of the PH TIC BN's performance.

6.6.1 **Diagnostic accuracy and factors associated with reduced performance**

Similar diagnostic accuracy statistics have been reported in other pre-hospital systems. In a sample of 433 trauma patients, clinicians from the Swiss helicopter emergency medical service missed 56% of abdominal injuries, 52% of pelvic injuries and 31% of chest injuries. The proportion of missed injuries was lower in head injuries (7%), femur fractures (10%) and tibia/fibula fractures (16%).[421] A study from SAMU in Paris demonstrates that difficulty in accessing the depth of penetrating injuries. In 153 patients with a penetrating thoracic injury pre-hospital physicians under triaged 42% and over triaged 30% of patients for the presence of an invasive thoracic wound.[422] Equally the difficulty in using pre-hospital clinical signs to diagnose raised intracranial pressure accurately has been studied. Closely mirroring the findings from this study, pre-hospital clinicians from AAKSS demonstrated relatively low sensitivity (48%) but high specificity (91%) in the diagnosis of raised ICP.[423]

Factors associated with missed injury are not well studied in the pre-hospital literature. The related research from in-hospital studies associates missed injuries with reduced consciousness,[424, 425] the severity of the patient's injuries,[426, 427] and the time of day.[428] Factors such as the clinician's experience, degree of fatigue, and presence of multiple patients, may be related to missed injuries. Further research is required to establish the relationship between these factors and missed injuries.

6.6.2 **Implication of variable uncertainty for Bayesian Network modelling**

Kyrimi and colleagues have developed a method to model the uncertain causal relationships present in medical Bayesian Networks.[429] They propose generally applicable and reusable medical reasoning patterns, named *medical idioms*. Idioms have a natural logic that mimics human reasoning process. Medical idioms represent essential reasoning steps in a patient's condition, diagnosis, prediction and management. The method extends to the uncertainty of diagnostic tests or patient reported symptoms, termed the *Manifestation reliability idiom*. The Manifestation reliability idiom explicitly describes how interpretation of signs, symptoms and tests are fallible. The effect is to accurately incorporate the clinical uncertainty pervasive in real world medical practice.

6.6.3 **Implication of reduced input certainty for model performance**

In the desire to develop decision support early after injury, the increased uncertainty of pre-hospital injury diagnosis is an important limiting factor. Differences exist between PH and ED access to effective imaging and POC investigations. Additionally, compensatory mechanisms make early clinical assessment challenging and occult injuries may present with overt physiological disturbance.[308] As a result, it appears pre-hospital injury diagnosis is more prone to error than ED diagnosis.[421]

Where previous ED modelling is accurate (see **1.9**) a reduction in the precision of the PH TIC BN's input variables will limit its performance. Thus, until reliable input variables can be entered into a pre-hospital model, it's performance will be inferior to in hospital modelling.

6.6.4 **Study limitations**

Several limitations are present in this study. The first limitation stems from the use of AIS scores. There are potential errors of transcription of AIS scores from clinical information into the databases. For example, in a study of pelvic injuries from the same institution (LAA), the pre-hospital sensitivity and specificity of pelvic fracture were better than the current study (0.69 and 0.81 respectively).[427] Notably, this study did not use AIS, but rather a binary classification of pelvic fracture: yes/no. In the present study, pre-hospital injuries were rarely given an AIS code of 3 or more, particularly with injuries to the abdomen. The reason for this is unclear. The assumed problem is pre-hospital underdiagnosis of the injury. However, an alternative explanation might be difficulty with the translation of the clinical injury diagnosis to the AIS scores. A lack of familiarity with the dictionary of scores, especially at the higher severity grades, maybe a source of measurement error in this study.

Another limitation is a result of the degree of specificity of the AIS used in this study. While each injury is given an AIS code, in the pre-hospital database, only the severity code is recorded (the post dot number). For example, AIS codes for a flail chest and haemothorax are 450209.3 and 442200.3. If a clinician suspected a flail chest and the

patient's final diagnosis was a haemothorax (without flail chest) this study would inappropriately classify the pre-hospital diagnosis as correct; as the pre-hospital and final injury scores have the same severity score of 3. The effect of this limitation is an overestimation of the ability of pre-hospital clinicians to diagnose injuries when only matching injury severity has been assessed.

Second, by analysis of patients only brought to the Royal London Hospital, it is uncertain whether a selection bias has been introduced, which may affect the generalisability of the findings to other MTCs. Third, the small sample size predisposes the analysis to statistical error. For this reason, minimal statistical analysis has been conducted on the data. This precludes useful techniques such as multivariable analysis to assess the factors associated with reduced clinical performance. A larger dataset would provide more confidence in such statistical analysis.

6.7 Conclusion

The pre-hospital clinical diagnosis of injuries of the chest, abdomen and pelvis is specific but not sensitive. Furthermore, severe injuries were not identified more accurately than less severe injuries. Reduced reliability of injury diagnosis pre-hospital, compared to in-hospital, has implications for pre-hospital models that rely on these inputs. Specifically, models that rely on clinical diagnosis can be expected to have inferior performance compared to models supported by accurate diagnostic adjuncts.

The next chapter systematically reviews the literature to ensure the causal relationships modelled by the BN represent the best structuring of knowledge, and consequently the best chance of accurate predictive performance.

CHAPTER 7

A SYSTEMATIC REVIEW OF THE CAUSES OF TRAUMA INDUCED COAGULOPATHY

7.1 Scope of the chapter

This chapter outlines the knowledge required for a Bayesian Network (BN) to predict trauma induced coagulopathy (TIC). The chapter systematically reviews the literature to determine the causes of trauma induced coagulopathy and assesses the evidence that supports these causal relationships. The evidence presented in each study is assessed using a modified form of Bradford Hill's criteria for causation. Each included study's risk of bias is quantified. The synthesised evidence is presented in tabulated and narrative forms.

7.2 Introduction

TIC is a post-injury disease state that encompasses disorders of coagulation and inflammation and is characterized by impairments in clot formation, breakdown, and overall haemostasis. TIC effects severely injured trauma patients and is associated with

poor patient outcomes and increased resource requirements. (see **1.4**) Advancing the understanding of TIC is therefore of utmost importance, and improvements in the management of TIC are likely to have substantial impact on reducing early trauma deaths due to haemorrhage and late deaths due to organ failure.

Fundamental to the practice of medicine, is the understanding of the causes that are necessary or sufficient to produce a disease. Through knowledge of these causal relationships, clinicians can make a range of useful inferences. For example, prediction of what will happen enables diagnosis and prognosis; controlling what happens enables targeted treatment; and noting what has happened enables construction of explanations.[430, 431]

However, distinguishing causal from non-causal associations requires careful appraisal of available experimental, empirical, and mechanistic evidence.[432] To assist this analysis, Hill's seminal paper in 1965 suggested several aspects of an association to consider when judging causation (Bradford Hill criteria).[433] Since publication Hill's criteria have been modified and updated.[434] These criteria provide a guide to systematically evaluate whether there is evidence to support a causal relationship. It is worth noting that these guidelines are not intended to be used simply as a checklist to establish causation but rather as a guide to critical appraisal of the evidence.

The aims of this review are to determine the causes of trauma induced coagulopathy and to identify, synthesise and assess the evidence that supports these causal relationships. This study is not intended to review the mechanistic causes of TIC, although where relevant to the discussion of the causes of TIC, the proposed mechanisms are briefly outlined. It is hoped that the review of the evidence will not just assist the development of future BN structures but also assist clinicians and researchers select patients at high risk of TIC.

7.3 Aims

7.3.1.1 *Aim 1:* Identify the causes of trauma induced coagulopathy

7.3.1.2 *Aim 2:* Assess the evidence that supports each causal relationships and each study's risk of bias

7.3.1.3 *Aim 3:* Synthesise the evidence

7.4 Methods

This systematic review was conducted in line with the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[435] The review protocol was registered with the International prospective register of systematic reviews (PROSPERO); registration number CRD42017057482.

7.4.1 Data sources and search strategy

Relevant publications were identified by an electronic search of the MEDLINE and EMBASE databases using a combination of the terms "trauma" and "coagulopathy". Searches were limited to English-language. The last search was performed on 01 January 2017. Two authors (M.M. and Z.P.) independently screened the search output for potentially relevant citations, and then assessed the full text of all identified citations for eligibility. Divergence was resolved by consensus with a third independent reviewer (R.D.). The reference lists of relevant articles were searched manually for additional relevant studies.

7.4.2 Study selection

Original studies that describe possible causes of TIC and provide evidence of the relationship with TIC, were included. A revised structure of Bradford Hill's criteria for causation was used to identify relevant evidence (**Table 28**).[433, 434] Studies were

excluded if they: i) did not clearly define the outcome of interest (TIC); ii) only reported methods to diagnose or treat TIC, with no description of potential causal mechanisms; iii) did not report original evidence; or iv) were published in abstract form only.

Table 28: A revised structure of Bradford Hill's criteria for causation

Criterion	Description
Strength (effect size)	A strong association is more likely to have a causal component than a modest association.
Consistency (reproducibility)	A relationship is observed repeatedly by different persons in different samples.
Specificity	A factor influences specifically a particular outcome or population with no other likely explanation.
Temporality	The effect occurs after the cause which may include a delay if there is an expected delay between the cause and expected effect.
Biological gradient	The outcome increases monotonically with increasing dose of exposure or according to a function predicted by a substantive theory.
Plausibility	A plausible mechanism between cause and effect is helpful.
Coherence	A causal conclusion should not fundamentally contradict present substantive knowledge, like laboratory findings.
Experiment	Causation is more likely if evidence is based on randomised experiments.
Analogy	For analogous exposures and outcomes an effect has already been shown.

Derived from references [433, 434]

7.4.3 **Risk of Bias assessment**

Two reviewers (M.M. and Z.P.) independently assessed the risk of bias in included studies. Studies were assessed according to predefined criteria (**Table 29**) using an adapted framework proposed by Altman[436] for observational and in-vitro studies, the Cochrane risk of bias tool[437] for randomised controlled trials, and an adapted version of SYRCLEs risk of bias tool for animal studies.[438] Each criterion was graded as low, unclear, or high risk of bias. Results between reviewers were compared and disagreements resolved by discussion. A summary score was calculated for each study by adding one point for each criterion graded as low risk of bias (score range is 0-7 for

randomised control trials, 0-9 for observational studies and 0-7 for experimental studies).

Table 29: Criteria for assessing Risk of Bias

Criteria
Observational studies, adapted from Altman[436]
Selection bias (Patient Sample)
1. Eligibility criteria
2. Baseline characteristics of sample
3. Representative sample from a relevant population
4. All eligible patients included
Measurement Bias (Outcome)
5. Coagulopathy appropriately defined
6. Outcome assessed on admission
Measurement Bias (Causative Factor)
7. Causative factors defined
8. Causative factors measured accurately
Confounding
9. Adjustment for confounding
Randomised Controlled Trial, Cochrane risk of bias tool [437]
Selection Bias
1. Appropriate sequence generation
2. Allocation concealment
Performance bias
3. Blinding of participants and personnel
Detection bias
4. Blinding of outcome assessment
Attrition bias
5. Incomplete outcome data
Reporting bias
6. Selective reporting
Other bias
7. Other sources of bias
Animal studies, adapted from SYRCLEs risk of bias tool [438]
Selection Bias
1. Appropriate sequence generation
2. Baseline characteristics of sample described
3. Allocation concealment
Performance Bias
4. Blinding
Detection Bias
5. Blinding of outcome assessment
Attrition Bias
6. Incomplete outcome data
Other bias
7. Other sources of bias

7.4.4 **Data extraction**

Two reviewers (M.M. and Z.P.) independently extracted data using a standardized data collection form. The following predefined data was extracted from each study: i) study identifiers (first author, year of publication); ii) essential study data (study design, sample size, study population, country of origin); iii) coagulopathy definition; iv) potential causal factors investigated; v) evidence to support causality (Bradford Hill criteria).

7.4.5 **Qualitative evidence synthesis**

Information from each study was stratified according to the potential causal factors of TIC identified. Studies within each causal group were then sub-grouped according to study design (Human observational, human experimental or animal experimental). The evidence for a causal relationship between each factor and TIC was then systematically evaluated according to a modification of Bradford Hill's criteria for causation (**Table 28**).[433, 434] Results from each study were tabulated and summarised in a narrative. A quantitative summary of results was not possible due to the methodological and clinical heterogeneity between included studies.

7.5 Results

7.5.1 **Search Results**

The search identified 3222 unique citations, of which 427 were potentially relevant. Overall, 168 studies were included in the evidence synthesis (**Figure 37**). One hundred ninety three studies were excluded because of an ineligible study type: abstract only (99), not original studies (94). A further, 66 studies were excluded as they did not measure coagulation (15) or did not describe a cause of coagulopathy (51).

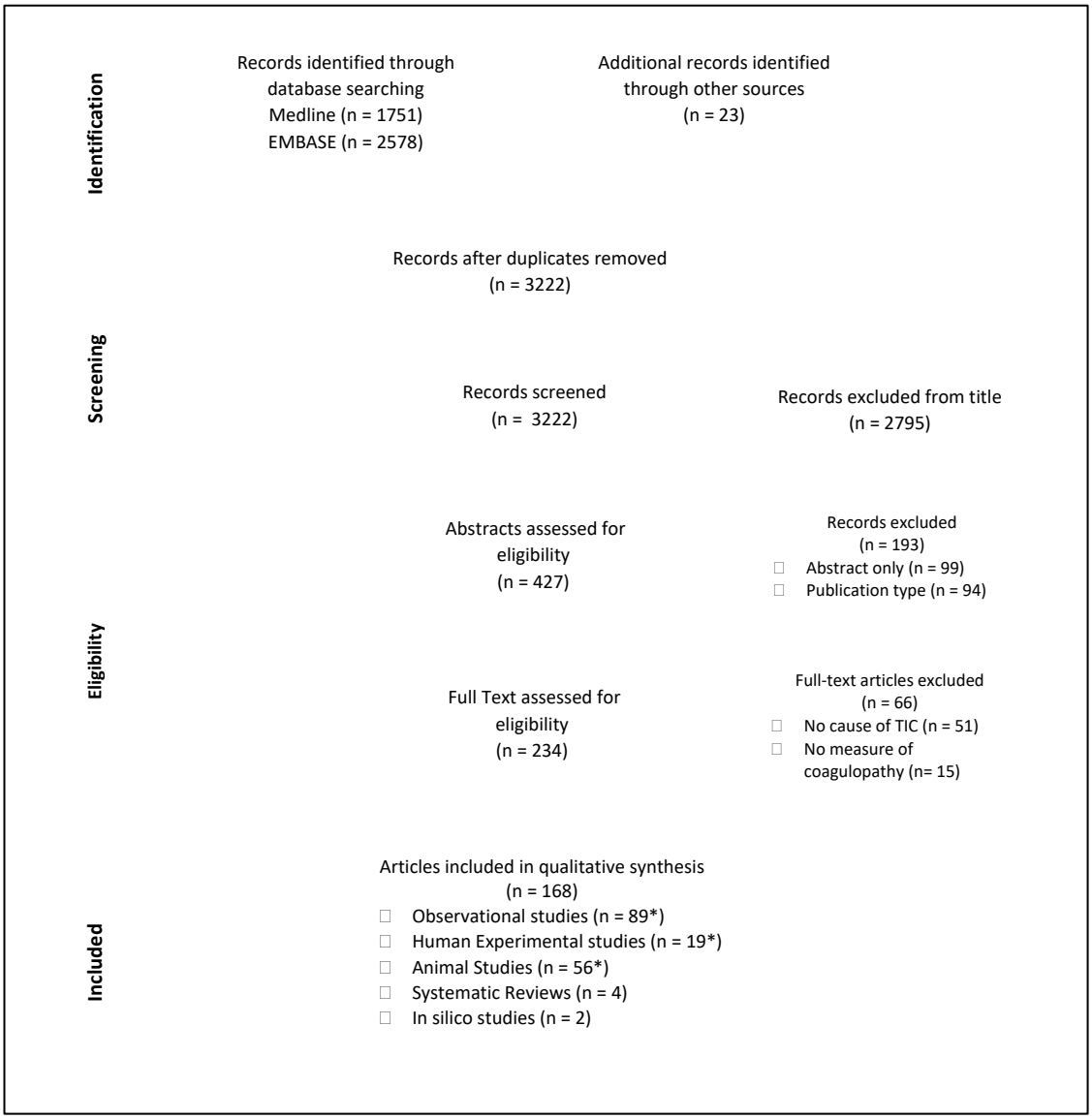

Figure 37: PRISMA flow diagram

7.5.2 **Study characteristics**

Eighty-nine human observational studies (50 prospective, 36 retrospective, 3 unclear), 19 human experimental studies and 56 animal studies were included. Two articles describe two distinct studies each. In addition, we included evidence from four systematic reviews and two in-silico modelling studies (**Figure 37**). Included studies originated from 22 countries, with approximately half published within the last five years of the search period (mode 2013, range 1974-2016) and nearly half (78/163) of the studies originate from US research institutions.

Human observational studies originated from 20 countries, and selected trauma patients according to: age (44 studies), mechanism of injury (10 studies), time since injury (29 studies), injury severity (29 studies), specific injuries i.e. head injury (26 studies), admission physiology (26 studies), trauma system response (44 studies), treatment administered (24 studies), and survival time (11 studies). The most common exclusion criteria were a known bleeding diathesis or pre-injury anticoagulation treatment (48 studies). The average sample size of human observational studies was 150 (range: 11 to 8724) patients. The average incidence of TIC in study populations was 0.24 (Range: 0.03 to 0.65)

Included human experimental studies comprised twelve in-vitro blood studies, two in-vivo studies and two randomised control trials. The average sample size for in-vitro blood experiments was 11 (range: 6-45), in-vivo experiments 13 (range 9-16) and RCTs 26 (range 18-34) patients.

The majority of animal studies were randomised controlled trials (37 studies). Other study designs included: the development of animal models to investigate TIC (7 studies), non-randomised controlled experiments (6 studies), and exploratory experiments with no control group (4 studies). The primary aim of these animal studies was either to investigate the pathophysiology of TIC (35 studies), or to test a therapeutic (19 studies). These studies featured five different animal species. The majority were porcine (26 studies) or rat (22 studies) models, but mouse (2 studies), rabbit (2 studies), and sheep (2 studies) models were also used. The average sample size was 24 (range: 6 to 256) animals.

7.5.3 **Risk of bias**

The risk of bias assessment, including the grading of individual criteria and summary score, for each included study are tabulated in **Appendix C** and summarised below.

The majority of human observational studies had a low risk of bias in the following domains: representativeness of the patient sample (87/87, 100%), eligibility criteria reported (81/87, 93%), baseline characteristics described (69/87, 79%), coagulopathy defined (70/87, 80%), causative factors defined (72/87, 83%) and causative factors accurately measured (65/87, 75%). A high risk of bias was present in a number of studies because of the timely measurement of coagulopathy (50/87, 57%), inclusion of all eligible patients (36/87, 41%), and controlling confounding (7/87, 8%). The median summary score for human observational studies was 7 (IQR: 5.5-7) out of a maximum of 9.

The risk of bias of the human experimental studies varied by the type of study. The 12 in-vitro human blood experiments were well conducted with a low risk of bias; median summary risk of bias score 7 out of 7. The two in-vivo experiments had higher risk of bias both scoring 2 out of 7 and the RCTs scored 5 and 6 out of 7.

Most animal studies had a low risk of bias for appropriate sequence generation (37/44, 81%), baseline characteristics described (50/54, 93%), blind outcome assessment (41/44, 93%), incomplete outcome data (44/54, 82%), and other sources of bias (45/54, 83%). Of the 44 controlled animal studies, 34 (77%) had a high risk of bias because there was no blinding of investigators, and the reporting in 39 (89%) studies was unclear as to whether allocation to groups was adequately concealed. The median summary score for animal studies was 5 (IQR: 4-5) out of a maximum of 7.

7.5.4 **Definitions of TIC**

Human studies used multiple definitions of coagulopathy (**Table 30**). The most common coagulation measures used were: conventional coagulation tests such as International Normalised Ratio (INR), Prothrombin Time (PT), Partial Thromboplastin Time (PTT), activated Partial Thromboplastin Time (aPPT), Thrombin Time (TT) and Quick Value (80 studies); viscoelastic tests such as ROTEM, TEG, or Sonoclot (13 studies); clinical evidence of pathological bleeding (4 studies). The majority of studies used more than one coagulation measure to classify TIC (51/89, 57%) for example 8 studies used the DIC score to classify coagulopathic patients. Additionally, between studies, there was

variability in the threshold used to define abnormal coagulation for each coagulation measure. For example, INR was the most commonly used coagulation measure (49 studies), with seven different thresholds used to define TIC (range: >1.1 to >1.6).

Table 30: Methods used to classify coagulopathic patients in human observational studies

Tests used to classify coagulopathy	Number of studies	Thresholds
INR	49	> 1.1; >1.2; >1.3; >1.4; >1.5; >1.6; above normal
PT	29	>12.9; >13.3; >13.4; >13.5; >14.2; >14.6; >1; above reference range; >3-6s longer =1, >6s = 2 (DIC scoring); > 1.5x normal; x2 normal
APPT	27	>35; >36; >38; >40; >45; >59; >60; >1.2 x normal; > 1.5x normal
PTT	18	>29; >34; >34.4; >38; >38.4; >60; x2 normal; above reference range
Quick Value	9	Quick value <70% used by all studies
TEG	7	R>8; K>3; alpha <55; MA<51; LY30>3%
ROTEM	6	ExTEM CT >20% normal; ExTEM CFT >20% normal; ExTEM MCF <20% normal; ExTEM Ly30 > 1%/Cl30 < 95%; MCF FibTEM <20% normal; ExTEM CA5 <36mm; Max Lysis >15%; Max Lysis >20%
Other tests of coagulation	4	Tests included: Endogenous thrombin potential; Ethanol Gelation test; Consumptive Coagulopathy Grade; Euglobulin; bleeding time; Thrombin Time
Macroscopic evidence of coagulopathic bleeding	4	

APPT, Activated partial thromboplastin time; CFT, Clot formation time; CT, Clotting time; DIC, Disseminated intravascular coagulation; INR, International normalised ratio; LY30, Lysis at 30 minutes; MA, Maximum amplitude; MCF, Maximium clot firmness; PT, Prothrombin time; PPT, Partial prothrombin time; ROTEM, Rotational thromboelastometry; TEG, thromboelastography.

The in vitro, in vivo and animal experimental studies defined TIC as a significant change in coagulation measures compared to baseline measurements or a control group. Coagulation measures used in these experiments included conventional coagulation tests (48 studies); viscoelastic tests (51 studies); and bleeding time (8 studies).

7.5.5 Causal factors driving Trauma induced coagulopathy

Seven possible causes of TIC are described in the included studies: hypoperfusion, tissue injury, acidaemia, hypothermia, haemodilution, brain injury, and type of fluid infused. These factors frequently co-exist and are related to characteristics of the injury, degree of physiological derangement, and subsequent resuscitation interventions. Evidence supporting a causal relationship between the driving factor and TIC is presented in **Appendix D** (**Table 34 - Table 40**) and qualitatively described below.

7.5.5.1 Hypoperfusion

Systemic tissue hypoperfusion following trauma appears to be a principal cause of an early endogenous coagulopathy. The predominant cause of tissue hypoperfusion after trauma is haemorrhagic shock but tension pneumothorax, cardiac tamponade and central nervous system injury may occasionally contribute.[439] Observational studies of heterogeneous trauma populations, from a wide spectrum of geographic and socioeconomic settings, repeatedly demonstrate a strong correlation between systemic markers of tissue hypoperfusion and acute coagulation dysfunction. These markers include heart rate,[440] systolic blood pressure,[255, 257, 440-445] lactate,[444, 446] and base deficit.[54, 395, 440, 442, 444, 445, 447-459] Furthermore, experimental animal models that combine trauma and haemorrhagic shock are able to consistently induce an endogenous acute traumatic coagulopathy.[66, 395, 460-468] Many of these models carefully control for known exogenous confounders including hypothermia and haemodilution.[66, 395, 461, 464, 465, 468] In addition, endogenous coagulopathy predominately appears in the presence of tissue hypoperfusion[449-451, 469] and there is a clear dose-response relationship between the degree of tissue hypoperfusion and coagulation dysfunction.[395, 445, 447-449, 470] Although tissue hypoperfusion alone can cause coagulation dysfunction, this seems to be greatly amplified when combined with some degree of tissue injury.[395]

Tissue hypoperfusion appears to cause coagulopathy by activation of anticoagulant and fibrinolytic pathways.[54, 444, 445, 449, 451, 460, 470] The mechanisms for these processes are not fully understood. There is some debate on the relative importance of the effects of thrombomodulin on thrombin,[471] activation of protein C,[54, 440, 444-446, 449, 451, 472

endothelial release of tissue plasminogen activator,[473, 474] degradation of the endothelial glycocalyx,[81, 446, 475] relatively reduced plasminogen activator inhibitor,[476] dysregulated fibrin and fibrinogen metabolism,[188, 457, 477] platelet microparticles,[478] platelet dysfunction,[78, 79, 477, 479, 480] and concurrent activation of the immune system.[72, 82, 481] In parallel, coagulopathy is also commonly observed following other causes of systemic hypoperfusion such as severe sepsis,[482] severe burns,[483] and cardiac arrest.[484-487] The similarity of these findings strongly supports a causal hypothesis between systemic tissue hypoperfusion and coagulopathy (**Table 34**).

7.5.5.2 Tissue Injury

Tissue injury activates the clotting cascade by exposing tissue factor to blood and producing thrombin.[488, 489] By definition, all trauma patients have some degree of tissue injury. The extent depends on the mechanism of injury, the amount of energy transferred, and the proportion of the body involved. As no validated biomarker exists, the extent of tissue injury is commonly estimated using the Injury Severity Score (ISS). The ISS is an anatomical score that measures the overall severity of injured patients and was initially intended to predict mortality.[256] There is a significant association between coagulopathy and the extent of tissue injury, as measured by ISS.[54, 55, 255, 395, 442, 443, 445, 447, 453-455, 457, 471, 490-494] Furthermore, there appears to be a dose-response relationship between ISS and the proportion of patients that develop a coagulopathy.[55, 447, 493, 495] The components of the ISS that seem to contribute to this relationship are increasingly severe injuries to the head, chest, abdomen, and extremity AIS body regions, but not isolated face and external injuries.[55, 441, 442]

Coagulation dysfunction appears far more likely to develop, however, when tissue injury is combined with tissue hypoperfusion.[449, 450] Animal models support this observation: animals subjected to a combination of tissue injury and tissue hypoperfusion predictably develop an early coagulopathy, while coagulation remains normal in animals subjected to extra-cranial tissue injury alone.[395] Although ISS is designed to reflect the extent of tissue injury, it may also act as a marker of the volume of blood loss and shock. This is a potentially important confounder when using ISS as a measure of tissue injury.

There are conflicting reports on whether specific injuries are associated with an increased risk of coagulopathy. Major pelvic fractures[496, 497] and severe intra-abdominal haemorrhage[498, 499] have been associated with a high risk of developing coagulopathy. These injuries may be markers of overall tissue injury severity or simply risk factors for haemorrhage and shock. An alternative view is coagulopathy is independent of the dominant region of injury when AIS is used to describe injuries.[458] Thus, tissue injury appears to be an important initiator of coagulation and fibrinolysis, but alone, does not appear to cause coagulation dysfunction (**Table 35**).

7.5.5.3 Acidaemia

Acidaemia is defined as a blood pH less than 7.35. In injured patients, acidaemia is usually caused by lactic acidosis due to tissue hypoperfusion. Other potential causes include excess chloride administration, respiratory failure, and a lactic acidosis due to intense physical exertion.

In observational studies there is a strong association between acidaemia and coagulopathy in trauma patients.[255, 444, 500-503] In animal models, coagulopathy results from acidaemia (pH 7.1), caused by controlled haemorrhage and tissue hypoperfusion.[504, 505] To separate the effects of acidaemia and tissue hypoperfusion on coagulation function experimental models have administered exogenous acid. In such studies, exogenous acid leads to impairments in coagulation in both animal models[460, 465, 504, 506] and human blood.[507] In these experiments, worsening acidaemia impairs clotting function in a dose-dependent manner.[508, 509] Notably, simply reversing acidaemia does not appear to correct the coagulopathy on its own.[504-506, 510]
Mechanistic experiments suggest that the coagulation derangements produced by acidaemia are the result of increased fibrinogen consumption,[506, 511, 512] increased fibrinolysis,[508] and impaired coagulation protease and platelet function.[506, 509] (**Table 36**).

7.5.5.4 Hypothermia

Hypothermia is defined as a core body temperature of less than 35°C.[439] Following injury, hypothermia may be caused by an increase in heat loss resulting from environmental exposure, infusion of cold fluids, or the administration of anaesthetic drugs. In shocked patients, decreases in heat production may be due to reduced metabolism.

Observational studies have identified a significant association between hypothermia and coagulopathy in trauma patients.[255, 257, 440, 442, 500, 513] Using multivariable analysis, several studies have shown an independent association between hypothermia and TIC.[255, 257, 440, 502, 514, 515] Hypothermia seems to inhibit coagulation function in a dose-dependent[516-518] and reversible manner.[519]

Mechanistic studies suggest these changes may be the result of slowed enzyme activity,[514, 516-518] delayed thrombin generation,[512, 517, 520-523] decreased fibrinogen synthesis,[521] and altered platelet function.[514, 516, 518, 520] These mechanisms prolong clotting times but do not seem to affect the strength of formed clots.[464, 467, 520] However, clinically significant effects of hypothermia on coagulation function are only observed at temperatures below 33°C.[460, 518, 524] Above this temperature, mildly hypothermic trauma patients have similar coagulation function to patients with a normal body temperature.[451, 525, 526] (**Table 37**)

7.5.5.5 Dilution

In hypovolaemic shock, intracellular and extracellular fluid shifts into the vascular compartment to expand plasma volume. The fluid shifts result in haemodilution. This dilution is compounded by resuscitation with intravenous crystalloid and synthetic colloid fluids. TIC is associated with haemodilution as measured by admission haemoglobin concentration,[444] haematocrit,[492] and the volume of pre-hospital fluid administered.[440, 442, 454, 492, 527-529] There is a dose-dependent relationship between the degree of coagulation dysfunction and the volume of administered pre-hospital fluid.[444, 503, 528, 530]

The degree of shock and haemodilution, however, are closely correlated and it is again difficult to separate the clinical effects of tissue hypoperfusion from those of haemodilution on coagulation function. The independent coagulopathic effects of haemodilution have been demonstrated in in-vitro human blood,[464, 530-535] experimental animal models,[465, 536-540] and a computer simulation model.[541] As expected, administration of large volumes of intravenous fluid has a clear dilution effect on platelets and coagulation proteases. These mechanisms, in addition to a decrease in thrombin generation,[535-537, 542] decreased fibrinogen levels,[520, 536, 540, 543-545] increased fibrinolysis secondary to reduced antifibrinolytic factors,[531] and reduced platelet function,[535] prolong clot formation time and reduced clot strength. Experimental studies demonstrate that reversal of dilutional coagulopathy may be possible by administration of Fresh Frozen Plasma,[536, 537, 542] Cryoprecipitate,[535] fibrinogen,[535, 546, 547] Prothrombin Complex Concentrate,[536, 537, 546, 548] or recombinant Factor VIIa.[464, 537, 540, 549] (**Table 38**)

7.5.5.6 Brain injury

Experiments to induce coagulopathy have used brain tissue since the 1830's.[550] Patients with traumatic brain injury (TBI) and coagulopathy have a nine fold increase in mortality and are 35 times more likely to have a poor prognosis than TBI patients without coagulopathy.[551] However the causal relationship between isolated brain injury and coagulopathy remains an active area of research.[552] Human observational studies over four decades have shown an association between TBI and the development of an acute coagulopathy.[553-560] In a meta-analysis of 34 studies, one third of patients were coagulopathic after TBI.[551]

Contention arises over the specificity of brain induced coagulopathy. Many authorities state that coagulopathy after TBI relies on tissue hypoperfusion[450] and brain injury is no more a driver of ATC than any other tissue injury.[55, 458, 494, 552] Proponents of this theory argue, a catecholamine surge after brain injury produces systemic vasoconstriction leading to systemic hypo-perfused tissue which drives TIC.[561, 562] Another theory suggests when intracranial pressure rises, poor perfusion of the extensive capillary network within the brain parenchyma, may also drive a classical TIC. Conversely, some

data suggests the magnitude of traumatised brain is a more significant driver of coagulopathy than shock or hypoxia.[452, 495] Several authors report a dose-response relationship between the severity of brain injury, as measured by clinical indicators (e.g. Glasgow Coma Scale[441, 452, 456, 563-568]) and imaging techniques (e.g. CT[441, 456, 566, 568, 569]), and the development of a coagulopathy.

The mechanism of brain associated coagulopathy was classically ascribed to disseminated intravascular coagulation (DIC).[559, 570-572] DIC was believed to occur due to overwhelming activation of coagulation by exposure of tissue factor, which is especially prevalent on the adventitial surface of cerebral vessels. The result according to some authorities is a DIC phenotype with high levels of fibrinolysis and thrombin production early after injury.[83, 565, 573, 574] More recently platelet dysfunction,[575, 576] clotting factor depletion,[577] microparticles derived from blood,[578] and brain,[579, 580] cells and S100A10 driven fibrinolysis[581] have all been proposed as having distinct mechanistic roles.

Whether TBI induced coagulopathy is a separate pathophysiological process distinct from extra-cranial ATC is not clear. The literature contains mixed clinical evidence and limited experimental evidence. (**Table 39**)

7.5.5.7 Hypertonic Crystalloid and Colloid Resuscitation

In addition to the dilutional effect of fluid infusion on coagulopathy, certain specific fluids may exacerbate clotting dysfunction. Synthetic colloids (e.g. Dextran, Gelatine and Starch) and isotonic and hypertonic crystalloids have been shown in observational and experimental studies to have variable effects on clotting.[582] A human observational study including nearly 2000 patients between 2002-2007 noted a pre-hospital colloid:crystalloid ratio ≥1:2 ratio was independently associated with worse clotting function.[442] In recent years the use of synthetic colloids has all but gone from current resuscitation practice, due to a lack of evidence of survival benefit and potential harm associated with their use.[583-585]

The deleterious effects of synthetic colloids on clotting have been demonstrated over multiple molecular weights, formulations and dilutions.[464, 535, 542, 544, 586-589] Study designs which compare the same volume of added fluid between synthetic colloids and isotonic crystalloids demonstrate additional coagulopathy especially at higher dilutions with colloids.[590-592] For example in an RCT of 25 trauma patients, patients were randomised to receiving either a Gelatine colloid or an isotonic crystalloid. Patients in the colloid group had prolonged bleeding times.[589] The effect of synthetic colloids on clotting are greater than can be explained by pure dilution. Colloids with a greater molecular weight have more profound effects on coagulation.[593] Several plausible mechanisms have been suggested; inhibited platelet function by interaction with vWF[588] or blockade of the GIIb-IIIa fibrinogen receptor, impairment of fibrin polymerisation[594, 595] and increased systemic concentration of tPA due to reduced hepatic uptake.

Hypertonic crystalloids also exert an additional adverse effect on clotting. In a rare RCT, patients given a bolus of hypertonic saline had higher INRs on admission than patients given normal saline. In the hypertonic groups, the study found reductions in prothrombotic factors and increases in lytic factors.[596] In-vitro, VHA assays demonstrate prolonged clot formation time, reduced clot propagation and increased time to maximum clot strength with hypertonic saline.[597, 598] Amongst isotonic crystalloids, an RCT with 18 patients found a balanced crystalloid was associated with an increased rate of fibrin generation compared to 0.9% NaCl.[599] (**Table 40**)

7.6 Discussion

This systematic review identifies evidence that supports seven causes of TIC. Each causal factor is supported by evidence from more than one study design. This diversity supports Bradford Hill's similarity criteria of causality but makes quantitative summary of the evidence impossible. The identified literature spans over 40 years, 20 countries, six species and countless definitions of coagulopathy. However, only three small RCTs in humans were discovered.

This study is a synthesis of existing evidence and so the strength of the results is limited by what has been published previously. It is not possible to conclusively state that the casual factors identified in this review are definitely causes of trauma induced coagulopathy. We have presented evidence that *suggests* a causal relationship rather than proves the relationship.

Some of the causal factors are supported by a substantial quantity of studies. Despite the volume of evidence, and often considerable effort by the authors, many of the studies retain confounding between the casual factors. For example, in human observation or animal modelling it is difficult to identify the individual contributions to coagulopathy of tissue injury and haemorrhage.

This review has determined the causes of trauma induced coagulopathy by identifying, synthesising and assessing the published evidence that supports these causal relationships. In doing so the study clinicians and researchers are able to make a range of useful predictions in the diagnosis, prognosis and treatment of injured patients.

7.7 Conclusion

Multiple study designs provide casual evidence for a combination of endogenous and iatrogenic drivers of Trauma induced coagulopathy. There is evidence to support seven causes of coagulopathy: hypoperfusion, tissue injury, acidaemia, hypothermia, dilution, brain injury and choice of resuscitation fluid. Through understanding these causal relationships, clinicians can make useful inferences regarding diagnosis, prognosis and treatment. In addition, the knowledge identified in this chapter can be used to structure future Bayesian Networks to predict trauma induced coagulopathy.

CHAPTER 8

CONCLUSION

8.1 Summary of findings

This thesis contributes to the understanding of decision support after injury.

The thesis establishes the difficulty that pre-hospital clinicians face in the diagnosis of major haemorrhage and trauma induced coagulopathy, then demonstrates the potential value of decision support of a Bayesian Network model, before developing a novel Bayesian Network, which is subsequently assessed in a prospective impact study.

Chapter One describes the significance of bleeding and coagulopathy after injury and the improved patient outcomes associated with early intervention. A discussion of modern theories of decision making and the complexities of identifying high-risk patients illustrates the niche for decision support approaches. A critical appraisal of the existing decision support models describes the lack of impact studies and why no TIC and major haemorrhage models have been adopted into regular clinical practice. Attention is turned to the improvements in data collection and computing power that have enabled the development of effective machine learning techniques such as Bayesian Networks. The introduction chapter ends with an explanation of the strengths

of BNs and describes the impressive predictive performance of an existing Emergency Department TIC BN.

Chapter two identifies bleeding and coagulopathy after injury as challenges to decision making for pre-hospital clinicians. Pre-hospital clinicians use recognition-priming to make decisions about bleeding, transfusion and TIC and struggle to make decisions when faced with clinical uncertainty. Integrating probabilistic decision support into pre-hospital care presents opportunities (attention to essential cues and improved situational awareness) and potential challenges (threats to autonomy, difficulty handling probability, and clinician's apparent preference for gestalt).

Chapter three illustrates the potential clinical value of the existing Emergency Department BN TIC model for major haemorrhage prediction. Within the context of the London inclusive trauma system, the majority of major haemorrhage protocol activations are initiated pre-hospital. A proportion of these clinical decisions are made in error. When compared for accuracy, native clinical judgement and the BN had few statistical differences in their predictive performance. However, the ED TIC BN was more accurate than another decision support model; the ABC score.

Chapter four develops and validates a Bayesian Network model for the pre-hospital prediction of TIC and major haemorrhage. The PH TIC BN performance is only minimally degraded compared to the ED TIC BN and is more accurate than any other TIC model.

Chapter five prospectively assesses the impact of the novel PH TIC BN on clinical judgement in a prospective, multicentre study. The finding of no improvement in the performance of clinical judgement after decision support, may have been influenced by the study's design. Nevertheless, decision support was associated with small improvements in calibration, overall accuracy and ease of decision making. Notably, direct comparison of the model against expert clinical assessment was not different. This finding suggests the model may have utility supporting decision making in those providers with little medical training. The findings of this study will inform the design

of future randomised trials that aim to assess the model's impact on patient and health system outcomes.

Chapter six demonstrates inaccuracy in the pre-hospital clinical diagnosis of injuries. Clinical diagnosis is specific but not sensitive and the diagnosis of severe injuries is not found to be more reliable. The results in this chapter identify incorrect assumptions in Chapter 4 that led to over-inflated estimations of the prospective performance of the PH TIC BN observed in Chapter 5.

Chapter seven systematically reviews the existing literature to ensure the causal relationships modelled by the PH TIC BN represent the best structuring of knowledge, and consequently, the best chance of accurate predictive performance. Seven causes of coagulopathy are identified: hypoperfusion, tissue injury, acidaemia, hypothermia, dilution, brain injury and choice of resuscitation fluid. The precise understanding of these causal relationships enables optimal development of the structure of future Bayesian Networks to predict TIC.

8.2 Strengths and Limitations

Throughout the thesis, the strengths and limitations of each study were discussed. What follows are strengths and limitations that have not been discussed and are shared between multiple sections of the thesis.

8.2.1 **Strengths**

A key attribute of this thesis is the use of Bayesian Networks to compute risk predictions. BNs are flexible models that allow the incorporation of multiple pieces of information in a way that is causally structured. There are notable benefits of using BNs. For example, in some circumstances, causal models have better predictive performance than traditional machine learning approaches, which rely on associative inference.[600] But perhaps more importantly than small differences in predictive performance, BNs have readily explainable reasoning. Understanding the model's reasoning may enhance

the user's trust and increase the likelihood of the user incorporating the prediction into their decision making.[601] This is point is emphasised by the US National Artificial Intelligence Research and Development strategic plan: "trustworthy AI requires explainable AI".[602] By providing concise explanations that describe the most important chains of reasoning, clinicians may be more inclined to use BNs in their daily practice. However, we know like most healthcare prediction models, the majority of Bayesian Networks are never implemented in to regular clinical practice.[334] The reasons for this lack of adoption are touched on in Chapter 2 when exploring the challenges to introduction of a probabilistic decision support. Further specific investigation is required to understand the likely adoption and subsequent utility of a BN based CDST in pre-hospital trauma care.

Another strength of this thesis is the adherence to the reporting standards of the TRIPOD statement.[394] Multiple recent systematic reviews point to the incomplete reporting and imperfect study design of machine learning studies.[419, 603] This thesis reports metrics that are often neglected, including calibration, performance compared to clinicians, and evaluation of other prognostic models in the same datasets. Perhaps most importantly, this thesis has examined model performance in real-world conditions that bear direct relevance to routine clinical practice, in a way many AI studies do not.[604] The benefit of close adherence to the reporting guidelines includes the ability for meaningful critique and enhanced interpretation of the results.

Finally, this thesis has engaged end-users in the process of defining the initial problem requiring decision support, developing a novel model and then assessing the likely impact of the model on clinical judgement. It joins a small group of studies that prospectively evaluate models in real world cases. The results of this thesis informs the design of future studies to assess the impact of decision support in trauma on patient and health system outcomes.

8.2.2 **Limitations**

This thesis has moved towards an understanding of what impact a pre-hospital Bayesian Network model for the prediction of TIC and blood transfusion may have in practice.

However, despite real gains in knowledge, the impact of the model is not still fully understood. First, the thesis does not assess the impact of a complete CDST, but rather a model. The model is just one component of a CDST. Other elements include the interface between the model and the user. This interface must be easy to use and integrate well with the clinical workflow.[605] Impact studies must assess the impact of the *CDST* on important outcomes. Integration, model comprehensibility, and model use are critical determinants of model impact. The ease of use and perceived benefit cannot be assessed until the CDST is finalised.

Another limitation relates to how the model output is communicated to clinicians. The ultimate aim of decision support is to improve patient and health system outcomes. For changes in outcomes to be realised, a decision has to change. This thesis has mainly focused on changes in risk prediction with model information, rather than changes in decisions. The model's output in this thesis provides both a probability value and stratifies patients into risk categories. Providing the output in this manner is intended to assist decision making while dividing risk into useful strata for clinicians to base their decisions. There are potential problems with this assistive approach. For example, the end-user may not understand the clinical implication of each risk category. Alternatively, the model output can suggest decisions or courses of action for each risk category. A problem with this directive approach is it removes the ability of the user to weigh the costs of wrong decisions and requires the same cost function to be applied to all situations by all users.[606] This reduces the flexibility of the CDST and rather than augmenting decision making it forces decisions and erodes clinician's autonomy.

While the assistive approach allows for clinician judgment, a directive approach may have a greater effect on outcomes.[218] Studies of decision support in postoperative nausea and vomiting suggest that more significant patient benefit may be achieved when treatment recommendations are given to clinicians rather than isolated predictions.[274, 275, 607] Thus the advantage of prediction over classification may be lost in some applications. The form of the output (assistive vs directive) is an integral aspect of the final CDST. The model's output form, and the visual style in which that output is displayed to the user, have to be established before the impact of decision support can be definitively studied. This thesis does not consider these two options as mutually

exclusive. Future research should explore providing both probabilistic predictions and risk stratification for suggested therapeutic interventions.

The final limitation is the failure of the thesis to consider competing risks. The limited literature on pre-hospital decision support illustrates some of the unintended consequences of using decision support. Perhaps most importantly for pre-hospital trauma care is the risk that using a CDST may add additional pre-hospital time and increase patient harm. In a rare RCT of decision support, a CDST that aimed to improve appropriate referral to specialist resources in older patient falls was studied. The trial noted a two-fold increase in the referral rate, with no improvement in patient outcome and a 10% longer job cycle compared to the control arm.[608] Similar findings were noted in a simulated RCT. In this second RCT, nurses in the intervention arm were more adherent to pre-hospital guidelines but took twice as long as the control group.[609] Such competing risks need to be studied with patient and health system centred outcome metrics. Clinicians are unlikely to adopt a CDST into routine practice until the overall utility is established.

8.3 Avenues for future research

Several additional questions have been generated while investigating the hypotheses within this thesis. The first two questions relate to model design and the last three questions concern CDST impact analysis:

8.3.1 **Can the prediction of blood transfusion requirement be improved?**

The initial aim of predicting blood transfusion was simply to communicate risk within TIC risk categories. Blood transfusion, the need for damage control surgery and the risk of death are all readily understandable parameters that capture elements of a trauma patients' clinical trajectory. However, as the research evolved, the ability to predict blood transfusion requirement emerged as a potentially desirable objective in its own right. This was especially germane amongst civilian pre-hospital and emergency clinicians and military clinicians and planners.

To predict the *need* for blood transfusion, using a Bayesian Network approach such as *a treatment idiom*[429] requires casual reasoning. The outcome node "blood transfusion" could be modelled as child of hypovolaemia. Thus, the model would use the underlying state of hypovolaemia (due to blood loss) to predict transfusion requirement, rather than rely on the transfusion decisions made by clinicians on patients in a training dataset.

A future model may benefit from incorporating information gleaned from experts, as discussed in Chapter 2. For example, in Chapter 2 expert clinicians describe the cues they use to diagnose blood loss ("The Hateful Eight" Table 6). These signs could be included as evidence nodes for the hypovolaemia variable. The original parents of tissue perfusion (HR, SBP, Lactate and TIC) remain unchanged (**Figure 38**).

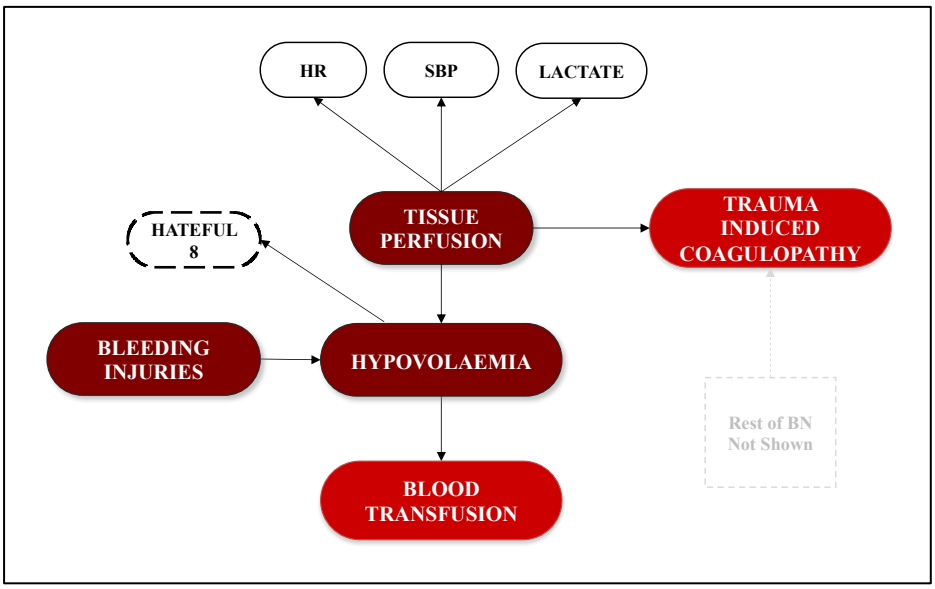

Figure 38: Bayesian Network fragment for blood transfusion prediction. Latent variables (dark red nodes) are causally related to the two outcome variables blood transfusion and coagulopathy (bright red nodes). The evidence measurements (white nodes) include "The Hateful Eight" (see Table 6) which have been grouped together for simplicity. HR, heart rate; SBP, systolic blood pressure;

Future research should explore the predictive performance of a BN that can give predictions at several times after injury. The ability to accurately predict the need for both pre-hospital blood transfusion and early in-hospital blood transfusion would be beneficial to multiple stakeholders. Using dynamic BNs, a single BN model could

generate predictions, that are updated, using data as it became available during the patient journey.

A more ambitious modelling goal, identified by the end-users, is to predict the cause of shock after injury. Clinicians described uncertainty in their ability to differentiate between types of shock. Specifically, differentiating between haemorrhagic shock, that should be treated with blood transfusion, and "haemorrhage mimics", i.e. all the other types of shock that do not benefit from transfusion. This goal is ambitious as it requires an appropriately labelled dataset to learn the relationships between variables.

8.3.2 Will inclusion of new variables improve the predictive performance of the models?

In Chapter 5 clinicians demonstrated improved predictive performance at the end of the pre-hospital phase of care. This may be a result of clinicians updating their clinical judgement based on changes in vital signs or response to treatment. Updating the network with prognostically important variables should lead to an improvement in the model's predictive performance. BNs can include dynamic variables (i.e. change over time) and produce predictions at more than one time point (**Figure 39**). The concept of adding new variables is straightforward. Once the causal structure for a target variable is defined, new evidence variables (e.g. the response dynamic response to blood transfusion) can be added to the causal chain.

In the future, variables that may improve performance could be identified from literature searches and discussion with experts. For example, there is literature supporting the diagnostic value of serial lactate measurements,[610] improvements in technology may detect blood loss more readily than traditional vital signs,[304] and interviews with expert clinicians reveal important cues that are helpful in the diagnosis of a hypotensive patient. However, the chief limitation in the inclusion of such candidate variables is the requirement for training datasets that contain these variables of interest.

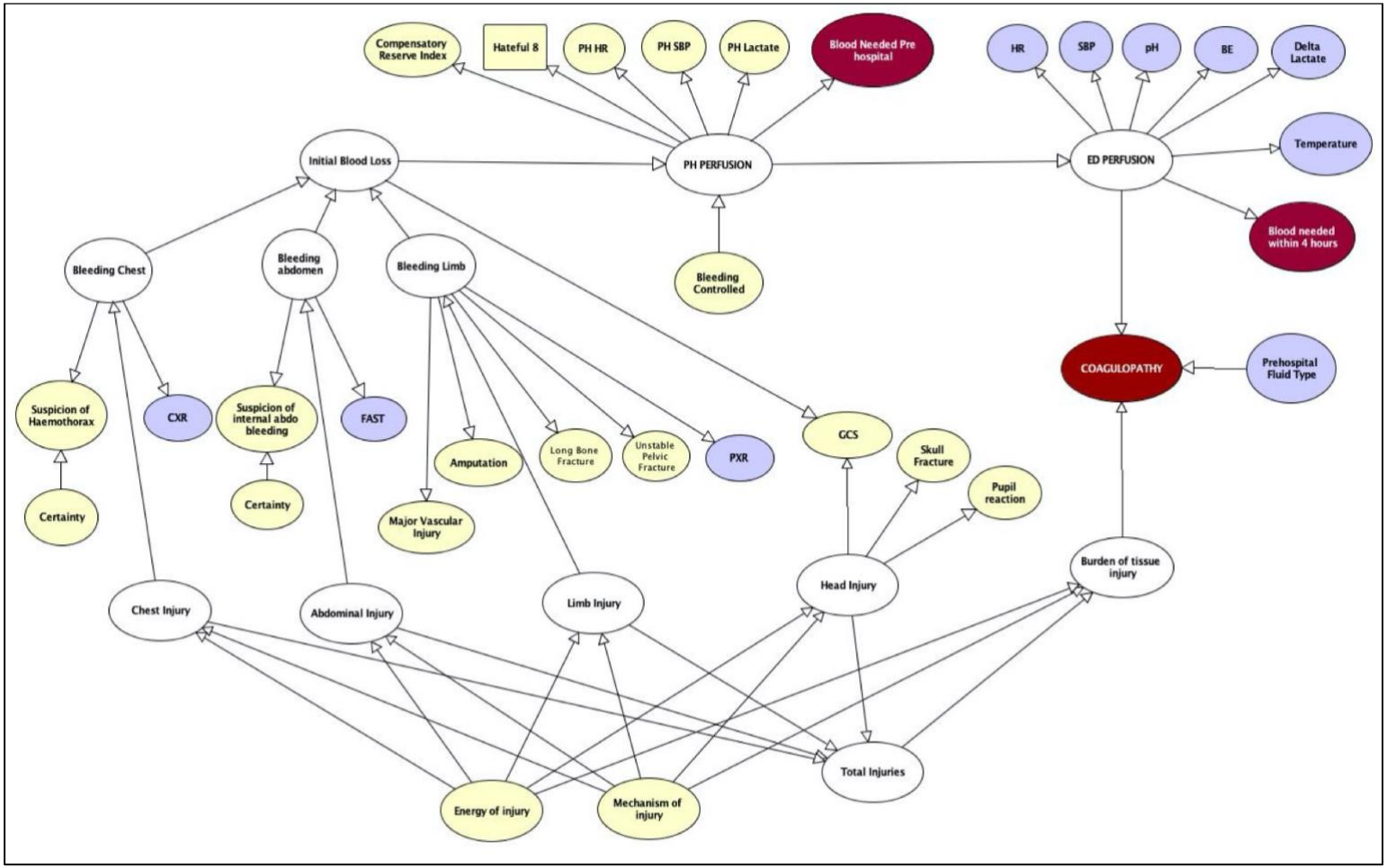

Figure 39: Directed Acyclic Graph of a future Bayesian Network developed to predict pre-hospital and early in-hospital blood transfusion, incorporating evidence from two distinct times and potential new evidence variables

Figure 39 (previous page) Directed Acyclic Graph of a future Bayesian Network developed to predict pre-hospital and early in-hospital blood transfusion, incorporating evidence from two distinct times and potential new evidence variables. The figure is produced in AgenaRisk software. The nodes represent outcome variables (red), evidence generated pre-hospital (yellow), evidence generated in hospital (blue), latent variables (white). The hateful eight have been grouped together for simplicity. HR, heart rate; SBP, systolic blood pressure; GCS, Glasgow coma scale; PH, Pre-hospital; ED, Emergency Department; CXR, chest x-ray; FAST; focused abdominal sonography in trauma; SBP, systolic blood pressure; BE, base excess.

The proposed BN model in **Figure 39** retains the basic causal structure as the previous models used in this thesis. Notably this model has two outputs: TIC and blood transfusion and two-time points PH and ED. The model has been updated by using new variables understood to be useful over the course of this thesis: first, it contains new types of input variables such as the compensatory reserve index, the Hateful Eight, and treatments administered; second, it uses inputs from two distinct time points and could include changes in variables (e.g. the change in the lactate); third, it incorporates clinical uncertainty in the pre-hospital clinical diagnosis of injury; fourth, it increases the granularity of injury descriptions to improve the description of total body injury; finally, the causal structure has been updated to include the causal factors identified in the systematic review in Chapter 7. Further modifications, such as modeling the need for blood transfusion as in **Figure 38** may improve the predictive performance of the model. Training and assessing the performance of this model requires further research and development.

8.3.3 How can the CDST interface maximise ease of use to optimise adoption?

The Technology Acceptance Model (TAM) states that innovations that are perceived as having value, and are easy to use, produce user attitudes that lead to intention to use the innovation and actual use.[611] Much of the work in this thesis is about establishing the model's value. The thesis does not address how to make the model easy to use. From the TAM statement it is clear that the ease of use is vital to a CDST. This is especially true in the time pressured context of pre-hospital trauma care. In this high stakes environment, the interface between the end-user and the model is critical for the successful integration of decision support into routine care.

A core requirement for a CDST is that it integrates data-handling and user interfaces with model computation and explanation. Data capture should be as unobtrusive as possible. At present, the data that powers the current models must be entered manually by a user, directly interacting with the interface. Short term goals that should be addressed include data entry that is optimised for a touch screen on ruggedised tablet platforms. On this platform users should interact with simple diagrams of anatomic injuries to enable facile data capture. In the medium term, the interface should leverage increasingly common technologies such as natural language processing and wireless data transfer from electronic devices such as heart rate monitors. Longer-term goals may include body-worn cameras to capture data such as pallor, HR, SBP or injury data that could be transferred to the model for predictions to be performed without any human interaction with the interface.

While many of these objectives are engineering tasks that require regular interaction between stakeholders, important research questions remain relating to usability. These questions need prospective study designs to address them. Similarly, research efforts to address how best to communicate the model's output to the user requires careful attention to lessons from behavioural psychology and prospective study to quantify clinical understanding.

8.3.4 What is the impact on patient and health system outcomes of predicting TIC and blood transfusion?

The performance of the PH TIC BN model in this thesis meets that exhibited by specialist clinicians. However, there are unanswered questions about whether the model can augment clinical decision making and impact important outcomes. Definitive impact analysis on patient and health system centred outcomes requires a finalised CDST, which is not yet available. Meanwhile, the research questions suggested above may improve predictive performance of a model. Thus, the timings of model finalization, completion of CDST development and trialing via RCT needs to be staged and synchronised. Moving to an RCT with a suboptimal CDST wastes time and resources. On the other hand, needlessly iterating the CDST before clinical impact

studies are considered will delay a potentially useful model. Undoubtedly, a phase III impact study is required to assess the impact of decision support, but the criteria to identify precisely when to start that trial and what constitutes an acceptable level of optimization remains unclear.

8.3.5 **What governs the outcome when AI and Humans disagree?**

Artificial Intelligence promises to analyse enormous quantities of data, identify patterns and provide the user with outputs that are impossible to attain by humans alone. As such, there is little doubt that AI will increasingly participate in medical decision making.[279] In an attempt to understand what future relationships with AI will take, Neves and Marsh described three simplified uses of AI in healthcare: AI replaces the human, AI filters for the human, and AI assists the human.[612] The models depicted in this thesis fall into later of the three uses. Our AI aims to assist and augment human decision making. Taking a broad view of medical AI, Topol argues that humans should remain "in the loop" and that algorithms in medicine should not replace humans.[310]

The assistance paradigm of AI requires a human-machine pairing. In this pairing, each agent (human and AI) may take one of several states. Although not directly demonstrated in the results of the experimental chapters, during the research two pertinent observations to AI-human pairing were noted. First, it was noted that when the AI - human pair agree, the behaviour may not change; as the clinician's decision remains unchanged. Arguably, this pairing state in which both agents agree is the lens through which much of the promise of medical AI is viewed.

The second observation pertains to what happens when the clinician disagrees with the decision support output. When the agents are discordant, the extent to which the clinician will allow the tool to modify behaviour appears to depend on several factors. Understanding what factors influence the decision outcome in agent discordance, sits within a broader theme. The wider research question encompasses how to integrate AI approaches into healthcare effectively.

Central to this issue is the need for the machine to be able to explain its reasoning, so-called explainable AI. The AI-human pairing should be re-considered as a partnership. When both agents hold the same "view", decision making is likely to be straightforward. But when the agents are discordant, perhaps the most critical factor that influences the outcome will be the machine's ability to explain. The machine will need to have clear explanations that are easily interpretable to *persuade* the human agent to act differently. And it's not just that single decision the human's trust in the machine's explanation is likely to influence longer-term device adoption as well.

What constitutes an effective explanation? Elements of explanation through visualisation have been explored in this thesis and are addressed elsewhere.[613] However, visualisations incompletely address effective explanation. If the aim is to persuade and change the perspective of a human agent, what is required is a diplomatic articulation of the explanation. Persuasion is a skill that all humans can relate to. A range of ideas come to mind when we consider what techniques we might employ to persuade a colleague to align with our position. Mostly these techniques require a conversation or at least interaction between the agents. Drawing on this human experience of explanation and persuasion is where others argue that explainable AI should learn from. There is much to learn from the fields of philosophy, psychology, and cognitive science. Rich research over hundreds of years describes how humans define, generate, select, evaluate, and present explanations. These fields also discuss the associated cognitive biases and social expectations of explainations.[614] The application of this rich theoretical work, to our specific decision support problem and the broader AI-human partnership requires further research.

Of course, explanation and interpretability are not the only factors that govern the outcome of agent discordance. For example, there are suggestions, raised in this thesis, that the degree to which a clinician will change her decision also varies between individuals and contexts. A future research aim is, therefore, to map the factors that influence whether the human will reverse their decision given new evidence. Once the elements are better understood, attempts can be made to optimise them in favour of the human and AI partnership better co-operating.

This argument is predicated on the assumption that the AI performs better on average than the human. If the AI does not perform as well, or systematically performs less well in certain circumstances (e.g. a specific injury pattern), then persuasion of the human agent may cause harm. In this situation, analysis of the cause of divergence between clinicians and the AI could generate knowledge that can be used to inform model re-design; this is especially germane in the application of Bayesian Networks.

8.4 The next steps

We plan further research to embed accurate AI risk prediction algorithms into a usable and useful clinical decision support systems to improve situational awareness and clinical decision-making. Since submission of this thesis, in November 2020, further work has been conducted to understand of the implementation stages. A scoping review has been performed to identify the benefits, barriers and facilitating factors for implementing medical BNs in healthcare.[615, 616] Lessons learned from this literature review will be valuable in the implementation and clinical evaluation of the developed AI system. In addition, a scoping review of CDS adoption frameworks is being performed. The aim of this review is to explore the published literature and better understand the steps needed for translating an accurate AI model into a useful CDS system.

In the immediate future, to extend the work on evaluating the AI System in the pre-hospital environment we are planning a programme of work to evaluate a) AI system performance, b) user's performance, c) patient impact, d) safety of the AI-system, and e) gain sufficient evidence of system performance to enable application for regulatory approval. To achieve these aims a stepwise process will be used to implement and evaluate the AI system. This will be performed across three studies. During the first two studies, the AI system will undergo iterative modifications, informed by the evidence generated, to optimize system performance.

Study 1 will be similar to the EmPHATTIC study and will further examine implementation of an AI Risk Prediction and Decision Support System into a real-world

clinical setting in a prospective observational study. A step-wise process will be used to implement and evaluate the AI system. This will be performed in two phases, first an integration phase of the AI platform, followed by an early formative clinical evaluation phase of the AI system, using a prospective observational study to assess the system's usability, efficiency, interpretability, explainability, trust, potential impact and safety. One major difference to prior work will be the use of the model by the clinical team rather than a researcher. During these phases, the AI system and the way it is used will undergo iterative modifications, informed by the evidence generated, to optimize system performance. The study is a bridge between the in-silico development and simulation studies of the prior effort and the clinical trial of the AI system in the proposed effort (Study 3). The study will be reported according to the DECIDE-AI guideline.[617]

Study 2 examines usability and impact on clinical workflow of an AI Risk prediction system. The study will be a randomised, controlled simulation experiment with a crossover design. This study will move our research efforts into the military environment and will be conducted alongside the Royal Air Force (RAF) tactical medical wing's Medical Emergency Response Team (MERT) training course (see **1.3.5.1**). In both arms of the experiment, the MERT team will be presented with a standardised and realistic clinical simulation of an injured casualty with non-compressible haemorrhage. In the control simulation, the team will manage the case using standard care. In the intervention simulation, the team will manage the case using standard care augmented by the AI risk prediction system. Outcomes measured during the study will include key timings (time to administration of TXA, activation of the major haemorrhage protocol, initiation of blood transfusion), the number of errors, and the level of cognitive load participants experienced with key decisions. A human factors evaluation of the AI system will also be conducted in this study. This element of the study will assess human-system interaction and include a usability evaluation and an evaluation of the user's perception of the AI systems' interpretability, explainability, trustworthiness, and clinical value.

Study 3 will assess the whether the AI Risk Prediction System can augment a treating clinician's perception of an individual patient's risk of adverse outcomes (situational awareness) and subsequent clinical decisions, compared to standard unassisted clinician

performance. The study design is a pragmatic, controlled before-and-after trial performed across the London Trauma System. The study protocol will be reported according to the guidelines for clinical trial protocols for interventions involving artificial intelligence.[618] The study participants are the clinical decision-makers involved in the pre-hospital assessment and management of trauma patients. All clinicians working for LAA will be eligible for inclusion. In the intervention arm of the study, LAA clinicians will have access to the AI system in addition to standard care. The final version of the AI-system at the completion of Study 1 will be used in the trial with no further modifications of the system. The trial report will report the amount of missing input data and the effect on algorithm and clinician accuracy. The primary outcome measures are the accuracy of clinical judgement of the patient's risk of mortality, TIC and haemorrhage.

It is anticipated that this work will significantly advance the process of translating accurate AI prediction algorithms into useful and usable CDS systems so that the potential benefits of these digital healthcare technologies can be realised.

8.5 Conclusion

This thesis advances the understanding of the benefit and impact of pre-hospital decision support after injury. Engagement with the key users of the model has informed the development of a Bayesian Network for the prediction of blood transfusion and trauma induced coagulopathy in the pre-hospital arena. Unlike many studies, this research evaluated the impact of the model on clinical judgement in real world settings. The knowledge generated will be used to design trials that assess the impact of machine-learning powered decision support tools, on both individual patient outcomes and across trauma systems.

References

1. World Health Organisation. Injuries and violence: the facts 2014. Geneva; 2014.

2. Rhee P, Joseph B, Pandit V, et al. Increasing trauma deaths in the United States. Ann Surg 2014;260:13-21.

3. World Health Organisation. Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Available at: www.who.int/healthinfo/global_burden_disease/estimates/en/index2.html. Accessed 2/4/18.

4. Norton R, Kobusingye O. Injuries. N Engl J Med 2013;368:1723-30.

5. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine 2006;3:e442.

6. Patton GC, Coffey C, Sawyer SM, et al. Global patterns of mortality in young people: a systematic analysis of population health data. Lancet 2009;374:881-92.

7. Public Health England. Chapter 2: major causes of death and how they have changed. Health profile for England: 2017; 2017.

8. Moran CG, Lecky F, Bouamra O, et al. Changing the System - Major Trauma Patients and Their Outcomes in the NHS (England) 2008-17. EClinicalMedicine 2018;2-3:13-21.

9. Kehoe A, Smith JE, Edwards A, et al. The changing face of major trauma in the UK. Emerg Med J 2015;32:911-5.

10. Pallett JR, Sutherland E, Glucksman E, et al. A cross-sectional study of knife injuries at a London major trauma centre. Ann R Coll Surg Engl 2014;96:23-6.

11. Crewdson K, Lockey D, Weaver A, Davies GE. Is the prevalence of deliberate penetrating trauma increasing in London? Experiences of an urban pre-hospital trauma service. Injury 2009;40:560-3.

12. Office for National Statistics. The nature of violent crime in England and Wales: year ending March 2017; 2017.

13. Pearce AP, Marsden M, Newell N, et al. Trends in admission timing and mechanism of injury can be used to improve general surgical trauma training. Ann R Coll Surg Engl 2020;102:36-42.

14. ACS Committee on Trauma. Resources for Optimal Care of the Injured Patient. In: American College of Surgeons, ed. 6 ed. Chicago, IL 2014.

15. Nathens AB, Brunet FP, Maier RV. Development of trauma systems and effect on outcomes after injury. Lancet 2004;363:1794-801.

16. West JG, Trunkey DD, Lim RC. Systems of trauma care. A study of two counties. Arch Surg 1979;114:455-60.

17. Brown JB, Rosengart MR, Kahn JM, et al. Impact of Volume Change Over Time on Trauma Mortality in the United States. Ann Surg 2017;266:173-178.

18. Sewalt CA, Wiegers EJA, Venema E, et al. The volume-outcome relationship in severely injured patients: A systematic review and meta-analysis. J Trauma Acute Care Surg 2018;85:810-819.

19. A Joint Report from The Royal College of Surgeons of England and the British Orthopaedic Association. Better care for the severely injured; 2000.

20. National Confidential Enquiry into Patient Outcome and Death (NCEPOD). Trauma: Who cares?; 2007.

21. National Audit Office. Major trauma care in England; 2010.

22. Davenport R, Tai N, West A, et al. A major trauma centre is a specialty hospital not a hospital of specialties. Br J Surg 2010;97:109-117.

23. National Institute for Health and Care Excellence. Major trauma: service delivery [NG40]; 2016.

24. Ministry of Defence. Operational Patient Care Pathway - JSP 950; 2014.

25. Bricknell M. For debate: the Operational Patient Care Pathway. J R Army Med Corps 2014;160:64-9.

26. Greaves I. Military medicine in Iraq and Afghanistan: a comprehensive review. Boca Raton: CRC Press,; 2019. p. 1 online resource.

27. Penn-Barwell JG, Roberts SA, Midwinter MJ, Bishop JR. Improved survival in UK combat casualties from Iraq and Afghanistan: 2003-2012. J Trauma Acute Care Surg 2015;78:1014-20.

28. Sharpe D, McKinlay J, Jefferys S, Wright C. Military Prehospital Emergency Care: defining and professionalising the levels of care provided along the Operational Patient Care Pathway. J R Army Med Corps 2019;165:188-192.

29. Hodgetts TJ, Mahoney PF. Military pre-hospital care: why is it different? J R Army Med Corps 2009;155:4-8.

30. Hodgetts TJ. The future character of military medicine. J R Army Med Corps 2012;158:271-8.

31. Kragh JF, Jr., Walters TJ, Baer DG, et al. Survival with emergency tourniquet use to stop bleeding in major limb trauma. Ann Surg 2009;249:1-7.

32. Wandling MW, Nathens AB, Shapiro MB, Haut ER. Association of Prehospital Mode of Transport With Mortality in Penetrating Trauma: A Trauma System-Level Assessment of Private Vehicle Transportation vs Ground Emergency Medical Services. JAMA Surg 2018;153:107-113.

33. Karrison TG, Philip Schumm L, Kocherginsky M, et al. Effects of driving distance and transport time on mortality among Level I and II traumas occurring in a metropolitan area. J Trauma Acute Care Surg 2018;85:756-765.

34. Brown JB, Gestring ML, Guyette FX, et al. Helicopter transport improves survival following injury in the absence of a time-saving advantage. Surgery 2016;159:947-59.

35. Brown JB, Rosengart MR, Forsythe RM, et al. Not all prehospital time is equal: Influence of scene time on mortality. J Trauma Acute Care Surg 2016;81:93-100.

36. Holcomb JB. Transport Time and Preoperating Room Hemostatic Interventions Are Important: Improving Outcomes After Severe Truncal Injury. Crit Care Med 2018;46:447-453.

37. Cannon JW. Hemorrhagic Shock. N Engl J Med 2018;378:370-379.

38. Cothren CC, Moore EE, Hedegaard HB, Meng K. Epidemiology of urban trauma deaths: a comprehensive reassessment 10 years later. World J Surg 2007;31:1507-11.

39. Eastridge BJ, Mabry RL, Seguin P, et al. Death on the battlefield (2001-2011): implications for the future of combat casualty care. J Trauma Acute Care Surg 2012;73:S431-7.

40. Mitra B, Gabbe BJ, Kaukonen KM, et al. Long-term outcomes of patients receiving a massive transfusion after trauma. Shock 2014;42:307-12.

41. Halmin M, Chiesa F, Vasan SK, et al. Epidemiology of Massive Transfusion: A Binational Study From Sweden and Denmark. Critical care medicine 2016;44:468-77.

42. Rixen D, Siegel JH. Bench-to-bedside review: oxygen debt and its metabolic correlates as quantifiers of the severity of hemorrhagic and post-traumatic shock. Crit Care 2005;9:441-53.

43. Tisherman SA, Schmicker RH, Brasel KJ, et al. Detailed description of all deaths in both the shock and traumatic brain injury hypertonic saline trials of the Resuscitation Outcomes Consortium. Ann Surg 2015;261:586.

44. Champion HR, Lombardo LV, Wade CE, et al. Time and place of death from automobile crashes: Research endpoint implications. J Trauma Acute Care Surg 2016;81:420-6.

45. McKinley TO, McCarroll T, Metzger C, et al. Shock volume: Patient-specific cumulative hypoperfusion predicts organ dysfunction in a prospective cohort of multiply injured patients. J Trauma Acute Care Surg 2018;85:S84-S91.

46. Douzinas EE. Hemorrhagic shock resuscitation: a critical issue on the development of posttraumatic multiple organ failure. Crit Care Med 2012;40:1348-9.

47. Brohi KC, M. J.; Ganter, M. T.; Matthay, M. A.; Mackersie, R. C.; Pittet, J. F. Acute traumatic coagulopathy: initiated by hypoperfusion: modulated through the protein C pathway? Ann Surg 2007;245:812-8.

48. Davenport RA, Brohi K. Cause of trauma-induced coagulopathy. Curr Opin Anaesthesiol 2016;29:212-9.

49. Kornblith LZ, Moore HB, Cohen MJ. Trauma-induced coagulopathy: The past, present, and future. J Thromb Haemost 2019;17:852-862.

50. Dobson GP, Morris JL, Davenport LM, Letson HL. Traumatic-Induced Coagulopathy as a Systems Failure: A New Window into Hemostasis. Semin Thromb Hemost 2020;46:199-214.

51. Maegele M. Prediction models to advance individualized resuscitation in trauma hemorrhage and acute traumatic coagulopathy (ATC): even the longest journey starts with first steps-Lao-Tzu (Chinese philosopher). Ann Transl Med 2017;5:466.

52. Davenport RA, Guerreiro M, Frith D, et al. Activated Protein C Drives the Hyperfibrinolysis of Acute Traumatic Coagulopathy. Anesthesiology 2017;126:115-127.

53. Casstevens EC, MacLeod JB, Shaz B. Early trauma induced coagulopathy (ETIC) is associated with an increase in complications post-injury. Transfusion (Paris) 2010;50:36A-37A.

54. Cohen MJ, Call M, Nelson M, et al. Critical role of activated protein C in early coagulopathy and later organ failure, infection and death in trauma patients. Ann Surg 2012;255:379-85.

55. Brohi K, Singh J, Heron M, Coats T. Acute traumatic coagulopathy. J Trauma 2003;54:1127-30.

56. MacLeod JB, Lynn M, McKenney MG, et al. Early coagulopathy predicts mortality in trauma. J Trauma 2003;55:39-44.

57. Cole E, Gillespie S, Vulliamy P, et al. Multiple organ dysfunction after trauma. Br J Surg 2020;107:402-412.

58. Naumann DN, Hazeldine J, Midwinter MJ, et al. Poor microcirculatory flow dynamics are associated with endothelial cell damage and glycocalyx shedding after traumatic hemorrhagic shock. J Trauma Acute Care Surg 2018;84:81-88.

59. Johansson PI, Henriksen HH, Stensballe J, et al. Traumatic Endotheliopathy: A Prospective Observational Study of 424 Severely Injured Patients. Ann Surg 2017;265:597-603.

60. Naumann DN, Hazeldine J, Davies DJ, et al. Endotheliopathy of Trauma is an on-Scene Phenomenon, and is Associated with Multiple Organ Dysfunction Syndrome: A Prospective Observational Study. Shock 2018;49:420-428.

61. White NJ, Ward KR, Pati S, et al. Hemorrhagic blood failure: Oxygen debt, coagulopathy, and endothelial damage. J Trauma Acute Care Surg 2017;82:S41-S49.

62. Meledeo MA, Herzig MC, Bynum JA, et al. Acute traumatic coagulopathy: The elephant in a room of blind scientists. J Trauma Acute Care Surg 2017;82:S33-S40.

63. Brohi K, Cohen MJ, Ganter MT, et al. Acute traumatic coagulopathy: initiated by hypoperfusion: modulated through the protein C pathway? Ann Surg 2007;245:812-8.

64. Frith D, Stanworth S, Gaarder C, et al. Acute traumatic coagulopathy is dependent on both injury severity and shock: Clinical and experimental validation. J Thromb Haemost 2009;7:290.

65. Dobson GP, Letson HL, Sharma R, et al. Mechanisms of early trauma-induced coagulopathy: The clot thickens or not? J Trauma Acute Care Surg 2015;79:301-9.

66. Chesebro BB, Rahn P, Carles M, et al. Increase in activated protein C mediates acute traumatic coagulopathy in mice. Shock 2009;32:659-65.

67. Howard BM, Kornblith LZ, Cheung CK, et al. Inducing Acute Traumatic Coagulopathy In Vitro: The Effects of Activated Protein C on Healthy Human Whole Blood. PLoS One 2016;11:e0150930.

68. Kutcher ME, Kornblith LZ, Vilardi RF, et al. The natural history and effect of resuscitation ratio on coagulation after trauma: a prospective cohort study. Ann Surg 2014;260:1103-11.

69. Rizoli SB, Scarpelini S, Callum J, et al. Clotting factor deficiency in early trauma-associated coagulopathy. J Trauma 2011;71:S427-34.

70. Fries D, Martini WZ. Role of fibrinogen in trauma-induced coagulopathy. Br J Anaesth 2010;105:116-21.

71. Deguchi H, Sinha RK, Marchese P, et al. Prothrombotic skeletal muscle myosin directly enhances prothrombin activation by binding factors Xa and Va. Blood 2016;128:1870-1878.

72. Billiar TR, Vodovotz Y. Time for trauma immunology. PLoS Med 2017;14:e1002342.

73. Martini WZ. Fibrinogen metabolic responses to trauma. Scand J Trauma Resusc Emerg Med 2009;17:2.

74. Cotton BA, Harvin JA, Kostousouv V, et al. Hyperfibrinolysis at admission is an uncommon but highly lethal event associated with shock and prehospital fluid administration. J Trauma Acute Care Surg 2012;73:365-70; discussion 370.

75. Moore HB, Moore EE, Chin TL, et al. Activated clotting time of thrombelastography (T-ACT) predicts early postinjury blood component transfusion beyond plasma. Surgery 2014;156:564-9.

76. Cardenas JC, Wade CE, Cotton BA, et al. TEG Lysis Shutdown Represents Coagulopathy in Bleeding Trauma Patients: Analysis of the PROPPR Cohort. Shock 2019;51:273-283.

77. Jacoby RC, Owings JT, Holmes J, et al. Platelet activation and function after trauma. J Trauma 2001;51:639-47.

78. Solomon C, Traintinger S, Ziegler B, et al. Platelet function following trauma. A multiple electrode aggregometry study. Thromb Haemost 2011;106:322-30.

79. Kutcher ME, Redick BJ, McCreery RC, et al. Characterization of platelet dysfunction after trauma. J Trauma Acute Care Surg 2012;73:13-9.

80. Wohlauer MV, Moore EE, Thomas S, et al. Early platelet dysfunction: an unrecognized role in the acute coagulopathy of trauma. J Am Coll Surg 2012;214:739-46.

81. Ostrowski SR, Johansson PI. Endothelial glycocalyx degradation induces endogenous heparinization in patients with severe injury and early traumatic coagulopathy. J Trauma Acute Care Surg 2012;73:60-6.

82. Ostrowski SR, Henriksen HH, Stensballe J, et al. Sympathoadrenal activation and endotheliopathy are drivers of hypocoagulability and hyperfibrinolysis in trauma: A prospective observational study of 404 severely injured patients. J Trauma Acute Care Surg 2017;82:293-301.

83. Gando S. Disseminated intravascular coagulation in trauma patients. Semin Thromb Hemost 2001;27:585-92.

84. Gando S. Acute coagulopathy of trauma shock and coagulopathy of trauma: a rebuttal. You are now going down the wrong path. J Trauma 2009;67:381-3.

85. Gando S, Sawamura A, Hayakawa M. Trauma, shock, and disseminated intravascular coagulation: lessons from the classical literature. Ann Surg 2011;254:10-9.

86. Gando S, Wada H, Thachil J, et al. Differentiating disseminated intravascular coagulation (DIC) with the fibrinolytic phenotype from coagulopathy of trauma and acute coagulopathy of trauma-shock (COT/ACOTS). J Thromb Haemost 2013;11:826-35.

87. Gando S, Otomo Y. Local hemostasis, immunothrombosis, and systemic disseminated intravascular coagulation in trauma and traumatic shock. Crit Care 2015;19:72.

88. Gando S, Hayakawa M. Pathophysiology of Trauma-Induced Coagulopathy and Management of Critical Bleeding Requiring Massive Transfusion. Semin Thromb Hemost 2016;42:155-65.

89. Saver JL, Fonarow GC, Smith EE, et al. Time to treatment with intravenous tissue plasminogen activator and outcome from acute ischemic stroke. JAMA 2013;309:2480-8.

90. Ferrer R, Artigas A, Suarez D, et al. Effectiveness of treatments for severe sepsis: a prospective, multicenter, observational study. Am J Respir Crit Care Med 2009;180:861-6.

91. Clarke JR, Trooskin SZ, Doshi PJ, et al. Time to laparotomy for intra-abdominal bleeding from trauma does affect survival for delays up to 90 minutes. J Trauma 2002;52:420-5.

92. Kotwal RS, Howard JT, Orman JA, et al. The Effect of a Golden Hour Policy on the Morbidity and Mortality of Combat Casualties. JAMA Surg 2016;151:15-24.

93. Kotwal RS, Scott LLF, Janak JC, et al. The effect of prehospital transport time, injury severity, and blood transfusion on survival of US military casualties in Iraq. J Trauma Acute Care Surg 2018;85:S112-S121.

94. Meyer DE, Vincent LE, Fox EE, et al. Every minute counts: Time to delivery of initial massive transfusion cooler and its impact on mortality. J Trauma Acute Care Surg 2017;83:19-24.

95. Roberts DJ, Bobrovitz N, Zygun DA, et al. Indications for use of damage control surgery in civilian trauma patients. A content analysis and expert appropriateness rating study. Ann Surg 2016;263:1018-1027.

96. Cotton BA, Dossett LA, Au BK, et al. Room for (performance) improvement: provider-related factors associated with poor outcomes in massive transfusion. J Trauma 2009;67:1004-12.

97. Wafaisade A, Maegele M, Lefering R, et al. High plasma to red blood cell ratios are associated with lower mortality rates in patients receiving multiple transfusion (4</=red blood cell units<10) during acute trauma resuscitation. The Journal of trauma 2011;70:81-8; discussion 88-9.

98. Cotton BA, Reddy N, Hatch QM, et al. Damage control resuscitation is associated with a reduction in resuscitation volumes and improvement in survival in 390 damage control laparotomy patients. Ann Surg 2011;254:598-605.

99. Pommerening MJ, Goodman MD, Holcomb JB, et al. Clinical gestalt and the prediction of massive transfusion after trauma. Injury 2015;46:807-13.

100. American College of Surgeons. Committee on Trauma. Subcommittee on Advanced Trauma Life Support. Advanced trauma life support course for physicians: ACS; 1989.

101. Tchorz KM, International ATLS working group. Advanced trauma life support (ATLS®): the ninth edition. The Journal Of Trauma And Acute Care Surgery 2013;74:1363.

102. Guly HR, Bouamra O, Spiers M, et al. Vital signs and estimated blood loss in patients with major trauma: testing the validity of the ATLS classification of hypovolaemic shock. Resuscitation 2011;82:556-9.

103. Mutschler M, Nienaber U, Brockamp T, et al. A critical reappraisal of the ATLS classification of hypovolaemic shock: does it really reflect clinical reality? Resuscitation 2013;84:309-13.

104. Schultz W, McConachie I. Vital signs after haemorrhage – Caution is appropriate. Trends in Anaesthesia and Critical Care 2015;5:89-92.

105. Luginbühl M. Detection of haemorrhagic shock: Are vital signs obsolete? A commentary by Dr M. Luginbühl. Trends in Anaesthesia and Critical Care 2015;5:93-94.

106. Reisner AT, Edla S, Liu J, et al. Tachycardic and non-tachycardic responses in trauma patients with haemorrhagic injuries. Injury 2018;49:1654-1660.

107. Johnson MC, Alarhayem A, Convertino V, et al. Compensatory Reserve Index: Performance of A Novel Monitoring Technology to Identify the Bleeding Trauma Patient. Shock 2018;49:295-300.

108. Little RA, Kirkman E, Driscoll P, et al. Preventable deaths after injury: why are the traditional 'vital' signs poor indicators of blood loss? J Accid Emerg Med 1995;12:1-14.

109. Wo CC, Shoemaker WC, Appel PL, et al. Unreliability of blood pressure and heart rate to evaluate cardiac output in emergency resuscitation and critical illness. Critical care medicine 1993;21:218-23.

110. National Institute for Health and Care Excellence. Major trauma: assessment and initial management (NICE guideline 39). Available at: www.nice.org.uk/guidance/ng39.

111. Harris T, Davenport R, Mak M, Brohi K. The Evolving Science of Trauma Resuscitation. Emerg Med Clin North Am 2018;36:85-106.

112. Lendrum R, Perkins Z, Chana M, et al. Reply to: Prehospital REBOA: Time to clearly define the relevant indications. Resuscitation 2019;142:191-192.

113. Guyette FX, Meier EN, Newgard C, et al. A comparison of prehospital lactate and systolic blood pressure for predicting the need for resuscitative care in trauma transported by ground. J Trauma Acute Care Surg 2015;78:600-6.

114. Lewis CT, Naumann DN, Crombie N, Midwinter MJ. Prehospital point-of-care lactate following trauma: A systematic review. J Trauma Acute Care Surg 2016;81:748-55.

115. Ince C. Hemodynamic coherence and the rationale for monitoring the microcirculation. Critical care (London, England) 2015;19 Suppl 3:S8.

116. Naumann DN, Mellis C, Smith IM, et al. Safety and feasibility of sublingual microcirculation assessment in the emergency department for civilian and military patients with traumatic haemorrhagic shock: a prospective cohort study. BMJ open 2016;6:e014162.

117. Stewart CL, Mulligan J, Grudic GZ, et al. Detection of low-volume blood loss: compensatory reserve versus traditional vital signs. J Trauma Acute Care Surg 2014;77:892-7; discussion 897-8.

118. Moulton SL, Mulligan J, Grudic GZ, Convertino VA. Running on empty? The compensatory reserve index. J Trauma Acute Care Surg 2013;75:1053-9.

119. Hoffman M, Monroe III DM. A cell-based model of hemostasis. Thromb Haemost 2001;85:958-965.

120. Frith D, Brohi K. The acute coagulopathy of trauma shock: clinical relevance. Surgeon 2010;8:159-63.

121. Frith DG, J. C.; Gaarder, C.; Maegele, M.; Cohen, M. J.; Allard, S.; Johansson, P. I.; Stanworth, S.; Thiemermann, C.; Brohi, K. Definition and drivers of acute traumatic coagulopathy: clinical and experimental investigations. J Thromb Haemost 2010;8:1919-25.

122. Davenport R, Manson J, De'Ath H, et al. Functional definition and characterization of acute traumatic coagulopathy. Crit Care Med 2011;39:2652-8.

123. Mitra B, O'Reilly G, Cameron P. Validation of Point of Care INR measurement post trauma. EMA - Emergency Medicine Australasia 2012;24:22.

124. Holcomb JB, Minei KM, Scerbo ML, et al. Admission rapid thrombelastography can replace conventional coagulation tests in the emergency department: experience with 1974 consecutive trauma patients. Ann Surg 2012;256:476-86.

125. Khan S, Davenport R, Raza I, et al. Damage control resuscitation using blood component therapy in standard doses has a limited effect on coagulopathy during trauma hemorrhage. Intensive Care Med 2015;41:239-47.

126. Schochl H, Frietsch T, Pavelka M, Jambor C. Hyperfibrinolysis after major trauma: differential diagnosis of lysis patterns and prognostic value of thrombelastometry. J Trauma 2009;67:125-31.

127. Meledeo MA, Peltier GC, McIntosh CS, et al. Functional stability of the TEG 6s hemostasis analyzer under stress. J Trauma Acute Care Surg 2018;84:S83-S88.

128. National Institute for Health Research. Annual Report 2017-18; 2018.

129. Entegrion. Announces CE Mark for Point of Care System for Blood Coagulation Monitoring. Available at: https://www.entegrion.com/pcm-portable-coagulation-monitor/. Accessed 12/07/2019.

130. Hemosonics. FDA grants de novo for the HemoSonics' Quantra® QPlus® System. Available at: https://hemosonics.com/fda-grants-de-novo/. Accessed 12/07/2019.

131. Gill M. The TEG(R)6s on Shaky Ground? A Novel Assessment of the TEG(R)6s Performance under a Challenging Condition. J Extra Corpor Technol 2017;49:26-29.

132. Roberts TR, Jones JA, Choi JH, et al. Thromboelastography on-the-go: Evaluation of the TEG 6s device during ground and high-altitude Aeromedical Evacuation with extracorporeal life support. J Trauma Acute Care Surg 2019;87:S119-S127.

133. Mitra B, O'Reilly G, Collecutt M, et al. Prospective comparison of point-of-care international normalised ratio measurement versus plasma international normalised ratio for acute traumatic coagulopathy. Emerg Med Australas 2012;24:363-8.

134. Rotondo MF, Schwab CW, McGonigal MD, et al. 'Damage control': an approach for improved survival in exsanguinating penetrating abdominal injury. J Trauma 1993;35:375-82; discussion 382-3.

135. Rotondo MF, Zonies DH. The damage control sequence and underlying logic. The Surgical clinics of North America 1997;77:761-77.

136. Holcomb JB, Jenkins D, Rhee P, et al. Damage control resuscitation: directly addressing the early coagulopathy of trauma. J Trauma 2007;62:307-10.

137. Cannon JW, Khan MA, Raja AS, et al. Damage control resuscitation in patients with severe traumatic hemorrhage: A practice management guideline from the Eastern Association for the Surgery of Trauma. J Trauma Acute Care Surg 2017;82:605-617.

138. Lucas CE, Ledgerwood AM. Prospective evaluation of hemostatic techniques for liver injuries. J Trauma 1976;16:442-51.

139. Stone HH, Strom PR, Mullins RJ. Management of the major coagulopathy with onset during laparotomy. Ann Surg 1983;197:532-5.

140. Khan M, McMonagle M. Trauma: Code Red. 1st ed. Boca Raton: CRC Press; 2018.

141. Bickell WH, Wall MJ, Jr., Pepe PE, et al. Immediate versus delayed fluid resuscitation for hypotensive patients with penetrating torso injuries. N Engl J Med 1994;331:1105-9.

142. Sperry JL, Guyette FX, Brown JB, et al. Prehospital Plasma during Air Medical Transport in Trauma Patients at Risk for Hemorrhagic Shock. N Engl J Med 2018;379:315-326.

143. Howard JT, Kotwal RS, Stern CA, et al. Use of Combat Casualty Care Data to Assess the US Military Trauma System During the Afghanistan and Iraq Conflicts, 2001-2017. JAMA Surg 2019;154:600-608.

144. Robertson OH. Transfusion with Preserved Red Blood Cells. Br Med J 1918;1:691-5.

145. Robertson LB. Further Observations on the Results of Blood Transfusion in War Surgery: With Special Reference to the Results in Primary Hemorrhage. Ann Surg 1918;67:1-13.

146. Kiel F. Development of a blood program in Vietnam. Mil Med 1966;131:1469-82.

147. O'Reilly DJ, Morrison JJ, Jansen JO, et al. Initial UK experience of prehospital blood transfusion in combat casualties. J Trauma Acute Care Surg 2014;77:S66-70.

148. Weaver AE, Hunter-Dunn C, Lyon RM, et al. The effectiveness of a 'Code Red' transfusion request policy initiated by pre-hospital physicians. Injury 2016;47:3-6.

149. Smith IM, James RH, Dretzke J, Midwinter MJ. Prehospital Blood Product Resuscitation for Trauma: A Systematic Review. Shock (Augusta, Ga) 2016;46:3-16.

150. Morrison JJ, Oh J, DuBose JJ, et al. En-route care capability from point of injury impacts mortality after severe wartime injury. Ann Surg 2013;257:330-4.

151. O'Reilly DJ, Morrison JJ, Jansen JO, et al. Prehospital blood transfusion in the en route management of severe combat trauma: a matched cohort study. J Trauma Acute Care Surg 2014;77:S114-20.

152. Holcomb JB, Donathan DP, Cotton BA, et al. Prehospital Transfusion of Plasma and Red Blood Cells in Trauma Patients. Prehosp Emerg Care 2015;19:1-9.

153. Brown JB, Cohen MJ, Minei JP, et al. Pretrauma center red blood cell transfusion is associated with reduced mortality and coagulopathy in severely injured patients with blunt trauma. Annals of surgery 2015;261:997-1005.

154. Shackelford SA, Del Junco DJ, Powell-Dunford N, et al. Association of Prehospital Blood Product Transfusion During Medical Evacuation of Combat Casualties in Afghanistan With Acute and 30-Day Survival. JAMA 2017;318:1581-1591.

155. Holcomb JB, Swartz MD, DeSantis SM, et al. Multicenter Observational Prehospital Resuscitation on Helicopter Study (PROHS). Journal of Trauma and Acute Care Surgery 2017.

156. Rehn M, Weaver A, Brohi K, et al. Effect of Prehospital Red Blood Cell Transfusion on Mortality and Time of Death in Civilian Trauma Patients. Shock 2019;51:284-288.

157. Griggs JE, Jeyanathan J, Joy M, et al. Mortality of civilian patients with suspected traumatic haemorrhage receiving pre-hospital transfusion of packed red blood cells compared to pre-hospital crystalloid. Scand J Trauma Resusc Emerg Med 2018;26:100.

158. Brown JB, Guyette FX, Neal MD, et al. Taking the Blood Bank to the Field: The Design and Rationale of the Prehospital Air Medical Plasma (PAMPer) Trial. Prehospital emergency care : official journal of the National Association of EMS Physicians and the National Association of State EMS Directors 2015;19:343-50.

159. Chapman MP, Moore EE, Chin TL, et al. Combat: Initial Experience with a Randomized Clinical Trial of Plasma-Based Resuscitation in the Field for Traumatic Hemorrhagic Shock. Shock (Augusta, Ga) 2015;44 Suppl 1:63-70.

160. Moore HB, Moore EE, Chapman MP, et al. Plasma-first resuscitation to treat haemorrhagic shock during emergency ground transportation in an urban area: a randomised trial. Lancet 2018;392:283-291.

161. Pusateri AE, Moore EE, Moore HB, et al. Association of Prehospital Plasma Transfusion With Survival in Trauma Patients With Hemorrhagic Shock When Transport Times Are Longer Than 20 Minutes: A Post Hoc Analysis of the PAMPer and COMBAT Clinical Trials. JAMA Surg 2020;155:e195085.

162. Reitz KM, Moore HB, Guyette FX, et al. Prehospital plasma in injured patients is associated with survival principally in blunt injury: Results from two randomized prehospital plasma trials. J Trauma Acute Care Surg 2020;88:33-41.

163. Gruen DS, Brown JB, Guyette FX, et al. Prehospital plasma is associated with distinct biomarker expression following injury. JCI Insight 2020;5.

164. Sim ES, Guyette FX, Brown JB, et al. Massive transfusion and the response to prehospital plasma: It is all in how you define it. J Trauma Acute Care Surg 2020;89:43-50.

165. Smith IM, Crombie N, Bishop JR, et al. RePHILL: protocol for a randomised controlled trial of pre-hospital blood product resuscitation for trauma. Transfus Med 2018;28:346-356.

166. Borgman MA, Spinella PC, Perkins JG, et al. The ratio of blood products transfused affects mortality in patients receiving massive transfusions at a combat support hospital. J Trauma 2007;63:805-13.

167. Cotton BA, Gunter OL, Isbell J, et al. Damage control hematology: the impact of a trauma exsanguination protocol on survival and blood product utilization. J Trauma 2008;64:1177-82; discussion 1182-3.

168. Allcock EC, Woolley T, Doughty H, et al. The clinical outcome of UK military personnel who received a massive transfusion in Afghanistan during 2009. J R Army Med Corps 2011;157:365-9.

169. Pidcoke HF, Aden JK, Mora AG, et al. Ten-year analysis of transfusion in Operation Iraqi Freedom and Operation Enduring Freedom: increased plasma and platelet use correlates with improved survival. J Trauma Acute Care Surg 2012;73:S445-52.

170. Khan S, Allard S, Weaver A, et al. A major haemorrhage protocol improves the delivery of blood component therapy and reduces waste in trauma massive transfusion. Injury 2013;44:587-92.

171. Holcomb JB, Wade CE, Michalek JE, et al. Increased plasma and platelet to red blood cell ratios improves outcome in 466 massively transfused civilian trauma patients. Annals of surgery 2008;248:447-58.

172. Holcomb JB, Zarzabal LA, Michalek JE, et al. Increased platelet:RBC ratios are associated with improved survival after massive transfusion. J Trauma 2011;71:S318-28.

173. Stanworth SJ, Morris TP, Gaarder C, et al. Reappraising the concept of massive transfusion in trauma. Crit Care 2010;14:R239.

174. Marsden M, Benger J, Brohi K, et al. Coagulopathy, cryoprecipitate and CRYOSTAT-2: realising the potential of a nationwide trauma system for a national clinical trial. Br J Anaesth 2019;122:164-169.

175. Rossaint R, Bouillon B, Cerny V, et al. Management of bleeding following major trauma: an updated European guideline. Critical care (London, England) 2010;14:R52.

176. Roberts I, Shakur H, Coats T, et al. The CRASH-2 trial: a randomised controlled trial and economic evaluation of the effects of tranexamic acid on death, vascular occlusive events and transfusion requirement in bleeding trauma patients. Health Technol Assess 2013;17:1-79.

177. Morrison JJ, Dubose JJ, Rasmussen TE, Midwinter MJ. Military Application of Tranexamic Acid in Trauma Emergency Resuscitation (MATTERs) Study. Arch Surg 2012;147:113-9.

178. Morrison JJ, Ross JD, Dubose JJ, et al. Association of cryoprecipitate and tranexamic acid with improved survival following wartime injury: findings from the MATTERs II Study. JAMA Surg 2013;148:218-25.

179. Gayet-Ageron A, Prieto-Merino D, Ker K, et al. Effect of treatment delay on the effectiveness and safety of antifibrinolytics in acute severe haemorrhage: a meta-analysis of individual patient-level data from 40 138 bleeding patients. Lancet 2018;391:125-132.

180. Cole E, Davenport R, Willett K, Brohi K. Tranexamic acid use in severely injured civilian patients and the effects on outcomes: a prospective cohort study. Ann Surg 2015;261:390-4.

181. Jimenez JJ, Iribarren JL, Brouard M, et al. Safety and effectiveness of two treatment regimes with tranexamic acid to minimize inflammatory response in elective cardiopulmonary bypass patients: a randomized double-blind, dose-dependent, phase IV clinical trial. J Cardiothorac Surg 2011;6:138.

182. Cole E, Davenport R. Early tranexamic acid use in trauma haemorrhage: why do we give it and which patients benefit most? Int Emerg Nurs 2015;23:38-41.

183. collaborators C-, Roberts I, Shakur H, et al. The importance of early treatment with tranexamic acid in bleeding trauma patients: an exploratory analysis of the CRASH-2 randomised controlled trial. Lancet 2011;377:1096-101, 1101 e1-2.

184. Coats TJ, Fragoso-Iniguez M, Roberts I. Implementation of tranexamic acid for bleeding trauma patients: a longitudinal and cross-sectional study. Emerg Med J 2019;36:78-81.

185. Marsden MER, Rossetto A, Duffield CAB, et al. Prehospital tranexamic acid shortens the interval to administration by half in Major Trauma Networks: a service evaluation. Emerg Med J 2019;36:395-400.

186. Ker K, Roberts I, Shakur H, Coats TJ. Antifibrinolytic drugs for acute traumatic injury. Cochrane Database Syst Rev 2015:CD004896.

187. Myers SP, Kutcher ME, Rosengart MR, et al. Tranexamic acid administration is associated with an increased risk of posttraumatic venous thromboembolism. J Trauma Acute Care Surg 2019;86:20-27.

188. Moore HB, Moore EE, Gonzalez E, et al. Hyperfibrinolysis, physiologic fibrinolysis, and fibrinolysis shutdown: the spectrum of postinjury fibrinolysis and relevance to antifibrinolytic therapy. J Trauma Acute Care Surg 2014;77:811-7; discussion 817.

189. Kahneman D, Slovic SP, Slovic P, Tversky A. Judgment under uncertainty: Heuristics and biases: Cambridge university press; 1982.

190. Zsambok CE, Klein GA. Naturalistic decision making. Mahwah, N.J.: L. Erlbaum Associates; 1997.

191. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science 1974;185:1124-31.

192. Klein GA. The power of intuition: How to use your gut feelings to make better decisions at work: Crown Business; 2004.

193. Gladwell M. Blink : the power of thinking without thinking. 1st Back Bay trade pbk. ed. New York: Back Bay Books; 2007.

194. Klein G. Naturalistic decision making. Hum Factors 2008;50:456-60.

195. Klein GA, Orasanu J, Calderwood R, Zsambok CE. Decision making in action: Models and methods. Westport, CT, US: Ablex Publishing; 1993.

196. Schnittker R, Marshall S, Horberry T, et al. Exploring Decision Pathways in Challenging Airway Management Episodes. Journal of Cognitive Engineering and Decision Making 2017;11:353-370.

197. Johnson JG, Raab M. Take the first: Option-generation and resulting choices. Organ Behav Hum Decis Process 2003;91:215-229.

198. Argyris C. Double loop learning in organizations. Harv Bus Rev 1977;55:115-125.

199. Meehl PE. Clinical versus statistical prediction; a theoretical analysis and a review of the evidence. Minneapolis,: University of Minnesota Press; 1954.

200. Kahneman D, Klein G. Conditions for intuitive expertise: a failure to disagree. Am Psychol 2009;64:515-26.

201. Sox HC, Blatt MA, Higgins MC, Marton KI. Medical decision making: ACP Press; 2007.

202. Donaldson MS, Corrigan JM, Kohn LT. To err is human: building a safer health system: National Academies Press; 2000.

203. Knight FH. Risk, uncertainty and profit. Boston and New York: Houghtin Mifflin Company 1921.

204. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? BMJ 2009;338:b375.

205. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10:e1001381.

206. Apgar V. A proposal for a new method of evaluation of the newborn. Classic Papers in Critical Care 1952;32:97.

207. Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. Br J Cancer 1982;45:361-6.

208. Agrawal A, Gans J, Goldfarb A. The simple economics of machine intelligence. Harv Bus Rev 2016;17.

209. Osheroff JA, Teich JM, Middleton B, et al. A roadmap for national action on clinical decision support. J Am Med Inform Assoc 2007;14:141-5.

210. Croskerry P. From mindless to mindful practice—cognitive bias and clinical decision making. New England Journal of Medicine 2013;368:2445-2448.

211. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. Circulation 2015;131:211-9.

212. Riley RD, van der Windt DA, Croft P, Moons KG. Prognosis Research in Healthcare: Oxford University Press; 2019.

213. Wyatt JC, Altman DG. Prognostic Models - Clinically Useful or Quickly Forgotten - Commentary. Br Med J 1995;311:1539-1541.

214. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ 2009;338:b604.

215. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-38.

216. Steyerberg EW, Harrell FE, Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol 2016;69:245-7.

217. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol 2015;68:25-34.

218. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med 2006;144:201-9.

219. El-Menyar A, Mekkodathil A, Abdelrahman H, et al. Review of Existing Scoring Systems for Massive Blood Transfusion in Trauma Patients: Where Do We Stand? Shock 2019;52:288-299.

220. Olaussen A, Thaveenthiran P, Fitzgerald MC, et al. Prediction of critical haemorrhage following trauma: A narrative review. Journal of Emergency Medicine, Trauma and Acute Care 2016;2016:3.

221. Nunez TC, Voskresensky IV, Dossett LA, et al. Early prediction of massive transfusion in trauma: simple as ABC (assessment of blood consumption)? J Trauma 2009;66:346-52.

222. Cotton BA, Dossett LA, Haut ER, et al. Multicenter validation of a simplified score to predict massive transfusion in trauma. J Trauma 2010;69 Suppl 1:S33-9.

223. American College of Surgeons - Committee on Trauma. Massive Transfusion in Trauma Guidelines; 2014.

224. Holcomb JB, Tilley BC, Baraniuk S, et al. Transfusion of plasma, platelets, and red blood cells in a 1:1:1 vs a 1:1:2 ratio and mortality in patients with severe trauma: the PROPPR randomized clinical trial. JAMA 2015;313:471-82.

225. Krumrei NJ, Park MS, Cotton BA, Zielinski MD. Comparison of massive blood transfusion predictive models in the rural setting. J Trauma Acute Care Surg 2012;72:211-5.

226. Mitra B, Rainer TH, Cameron PA. Predicting massive blood transfusion using clinical scores post-trauma. Vox Sang 2012;102:324-30.

227. Ogura T, Nakano M, Nakamura Y, et al. The Novel and Useful Scoring System to Predit Massive Transfusion for Trauma Patients: Traumatic Bleeding Severity Score (TBSS). Am Heart Assoc; 2012.

228. Brockamp T, Nienaber U, Mutschler M, et al. Predicting on-going hemorrhage and transfusion requirement after severe trauma: a validation of six scoring systems and algorithms on the TraumaRegister DGU®. Critical Care 2012;16.

229. Schroll R, Swift D, Tatum D, et al. Accuracy of shock index versus ABC score to predict need for massive transfusion in trauma patients. Injury 2018;49:15-19.

230. Motameni AT, Hodge RA, McKinley WI, et al. The use of ABC score in activation of massive transfusion. Journal of Trauma and Acute Care Surgery 2018:1.

231. Hodgman EI, Cripps MW, Mina MJ, et al. External validation of a smartphone app model to predict the need for massive transfusion using five different definitions. J Trauma Acute Care Surg 2018;84:397-402.

232. Hanna K, Harris C, Trust MD, et al. Multicenter Validation of the Revised Assessment of Bleeding and Transfusion (RABT) Score for Predicting Massive Transfusion. World J Surg 2020;44:1807-1816.

233. Motameni AT, Hodge RA, McKinley WI, et al. The use of ABC score in activation of massive transfusion: The yin and the yang. J Trauma Acute Care Surg 2018;85:298-302.

234. Heelan Gladden AA, Peltz ED, McIntyre RC, Jr., et al. Effect of Pre-Hospital Use of the Assessment of Blood Consumption Score and Pre-Thawed Fresh Frozen Plasma on Resuscitation and Trauma Mortality. J Am Coll Surg 2019;228:141-147.

235. Goodman MD, Hawes HG, Pommerening MJ. Prehospital ABC score accurately triages patients who will require immediate resource utilization. 72nd Annual Meeting of AAST and Clinical Congress of Acute Care Surgery; 2013 18-23 September 2013; San Francisco (CA), ; 2013.

236. Yucel N, Lefering R, Maegele M, et al. Trauma Associated Severe Hemorrhage (TASH)-Score: probability of mass transfusion as surrogate for life threatening hemorrhage after multiple trauma. J Trauma 2006;60:1228-36; discussion 1236-7.

237. Maegele M, Lefering R, Wafaisade A, et al. Revalidation and update of the TASH-Score: a scoring system to predict the probability for massive transfusion as a surrogate for life-threatening haemorrhage after severe injury. Vox Sang 2011;100:231-8.

238. Mina MJ, Winkler AM, Dente CJ. Let technology do the work: Improving prediction of massive transfusion with the aid of a smartphone application. J Trauma Acute Care Surg 2013;75:669-75.

239. Sun G-W, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. J Clin Epidemiol 1996;49:907-916.

240. Holcomb JB, del Junco DJ, Fox EE, et al. The prospective, observational, multicenter, major trauma transfusion (PROMMTT) study: comparative effectiveness of a time-varying treatment with competing risks. JAMA Surg 2013;148:127-36.

241. Olaussen A, Blackburn T, Mitra B, Fitzgerald M. Review article: shock index for prediction of critical bleeding post-trauma: a systematic review. Emerg Med Australas 2014;26:223-8.

242. Vandromme MJ, Griffin RL, Kerby JD, et al. Identifying risk for massive transfusion in the relatively normotensive patient: utility of the prehospital shock index. J Trauma 2011;70:384-8; discussion 388-90.

243. Mitra B, Fitzgerald M, Chan J. The utility of a shock index >/= 1 as an indication for pre-hospital oxygen carrier administration in major trauma. Injury 2014;45:61-5.

244. Olaussen A, Peterson EL, Mitra B, et al. Massive transfusion prediction with inclusion of the pre-hospital Shock Index. Injury 2015;46:822-6.

245. Terceros-Almanza LJ, Garcia-Fuentes C, Bermejo-Aznarez S, et al. Prediction of massive bleeding in a prehospital setting: validation of six scoring systems. Med Intensiva 2019;43:131-138.

246. Reed MJ, Glover A, Byrne L, et al. Experience of implementing a National pre-hospital Code Red bleeding protocol in Scotland. Injury 2017;48:41-46.

247. Hamada SR, Rosa A, Gauss T, et al. Development and validation of a pre-hospital "Red Flag" alert for activation of intra-hospital haemorrhage control response in blunt trauma. Critical care (London, England) 2018;22:113.

248. McLaughlin DF, Niles SE, Salinas J, et al. A predictive model for massive transfusion in combat casualty patients. J Trauma 2008;64:S57-63; discussion S63.

249. Larson CR, White CE, Spinella PC, et al. Association of shock, coagulopathy, and initial vital signs with massive transfusion in combat casualties. J Trauma 2010;69 Suppl 1:S26-32.

250. Schreiber MA, Perkins J, Kiraly L, et al. Early predictors of massive transfusion in combat casualties. J Am Coll Surg 2007;205:541-5.

251. McLennan JV, Mackway-Jones KC, Smith JE. Prediction of massive blood transfusion in battlefield trauma: Development and validation of the Military Acute Severe Haemorrhage (MASH) score. Injury 2018;49:184-190.

252. McLennan JV, Mackway-Jones KC, Horne ST, Body R. Predictors of massive blood transfusion: a Delphi Study to examine the views of experts. J R Army Med Corps 2017;163:259-265.

253. National Institute for Health and Care Excellence. Major trauma: Complex fractures. ; 2016.

254. Thorn S, Guting H, Maegele M, et al. Early Identification of Acute Traumatic Coagulopathy Using Clinical Prediction Tools: A Systematic Review. Medicina (Kaunas) 2019;55.

255. Cosgriff N, Moore EE, Sauaia A, et al. Predicting life-threatening coagulopathy in the massively transfused trauma patient: hypothermia and acidoses revisited. J Trauma 1997;42:857-61; discussion 861-2.

256. Baker SP, O'Neill B, Haddon W, Jr., Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma 1974;14:187-96.

257. Mitra B, Cameron PA, Mori A, et al. Early prediction of acute traumatic coagulopathy. Resuscitation 2011;82:1208-13.

258. Service QA. Clinical Practice Procedures: Assessment / COAST score; 2019.

259. Mitra B, Mazur S, Cameron PA, et al. Tranexamic acid for trauma: filling the 'GAP' in evidence. Emerg Med Australas 2014;26:194-7.

260. Thorn S, Tonglet M, Maegele M, et al. Validation of the COAST score for predicting acute traumatic coagulopathy: A retrospective single-centre cohort study. Trauma 2019;22:112-117.

261. Peltan ID, Rowhani-Rahbar A, Vande Vusse LK, et al. Development and validation of a prehospital prediction model for acute traumatic coagulopathy. Crit Care 2016;20:371.

262. Tonglet ML, Minon JM, Seidel L, et al. Prehospital identification of trauma patients with early acute coagulopathy and massive bleeding: results of a prospective non-interventional clinical trial evaluating the Trauma Induced Coagulopathy Clinical Score (TICCS). Crit Care 2014;18:648.

263. Tonglet M, Lefering R, Minon JM, et al. Prehospital identification of trauma patients requiring transfusion: results of a retrospective study evaluating the use of the trauma induced coagulopathy clinical score (TICCS) in 33,385 patients from the TraumaRegister DGU((R)). Acta Chir Belg 2017;117:385-390.

264. Swerts F, Mathonet PY, Ghuysen A, et al. Early identification of trauma patients in need for emergent transfusion: results of a single-center retrospective study evaluating three scoring systems. Eur J Trauma Emerg Surg 2019;45:681-686.

265. David JS, Voiglio EJ, Cesareo E, et al. Prehospital parameters can help to predict coagulopathy and massive transfusion in trauma patients. Vox Sang 2017;112:557-566.

266. Brohi K. Prediction of acute traumatic coagulopathy and massive transfusion - Is this the best we can do? Resuscitation 2011;82:1128-9.

267. Maegele M, Spinella PC, Schochl H. The acute coagulopathy of trauma: mechanisms and tools for risk stratification. Shock 2012;38:450-8.

268. Collins GS, Ogundimu EO, Cook JA, et al. Quantifying the impact of different approaches for handling continuous predictors on the performance of a prognostic model. Stat Med 2016;35:4124-35.

269. Janssen KJ, Vergouwe Y, Donders ART, et al. Dealing with missing predictor values when applying clinical prediction models. Clin Chem 2009;55:994-1001.

270. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.

271. Marsden MER, Mossadegh S, Marsh W, Tai N. Development of a major incident triage tool: the importance of evidence from implementation studies. BMJ Mil Health 2020;166:204.

272. Kappen TH, van Klei WA, van Wolfswinkel L, et al. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res 2018;2:11.

273. Sacks GD, Dawes AJ, Ettner SL, et al. Impact of a Risk Calculator on Risk Perception and Surgical Decision Making: A Randomized Trial. Ann Surg 2016;264:889-895.

274. Kappen TH, Moons KG, van Wolfswinkel L, et al. Impact of risk assessments on prophylactic antiemetic prescription and the incidence of postoperative nausea and vomiting: a cluster-randomized trial. Anesthesiology 2014;120:343-54.

275. Kappen TH, Vergouwe Y, van Wolfswinkel L, et al. Impact of adding therapeutic recommendations to risk assessments from a prediction model for postoperative nausea and vomiting. Br J Anaesth 2015;114:252-60.

276. Moore G. Cramming More Components onto Integrated Circuits. Electronics. 1965 19 April 1965:114-117.

277. Fry H. Hello world : being human in the age of algorithms. First edition. ed. New York, NY: W.W. Norton & Company; 2018.

278. Agrawal A, Gans J, Goldfarb A. Prediction machines : the simple economics of artificial intelligence. Boston, Massachusetts: Harvard Business Review Press; 2018.

279. Topol EJ. The Topol Review: Preparing the healthcare workforce to deliver the digital future: Health Education England,; 2019.

280. Turing A. Computing machinery and intelligence. Mind 1950;59:433.

281. The Royal Society. Machine learning: the power and promise of computers that learn by example; 2017.

282. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods 2018;15:233-234.

283. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision 2015;115:211-252.

284. Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning. arXiv preprint arXiv:13125602 2013.

285. Witten IH, Frank E, Hall MA, Pal CJ. Data mining : practical machine learning tools and techniques. Fourth Edition. ed. Amsterdam: Elsevier; 2017.

286. Singh R, Kalra MK, Nitiwarangkul C, et al. Deep learning in chest radiography: Detection of findings and presence of change. PLoS One 2018;13:e0204155.

287. Nam JG, Park S, Hwang EJ, et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. Radiology 2019;290:218-228.

288. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. JAMA 2017;318:2199-2210.

289. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng 2018;2:158-164.

290. Abramoff MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018;1:39.

291. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115-118.

292. Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018;29:1836-1842.

293. Mori Y, Kudo S-e, Misawa M, et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. Ann Intern Med 2018;169:357-366.

294. Zhang J, Gajjala S, Agrawal P, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. Circulation 2018;138:1623-1635.

295. Wang X, Peng Y, Lu L, et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017: IEEE; 2017. p. 3462-3471.

296. Lindsey R, Daluiski A, Chopra S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci U S A 2018;115:11591-11596.

297. Titano JJ, Badgeley M, Schefflein J, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. Nat Med 2018;24:1337-1341.

298. Liu NT, Salinas J. Machine Learning for Predicting Outcomes in Trauma. Shock 2017;48:504-510.

299. Walczak S. Artificial neural network medical decision support tool: predicting transfusion requirements of ER patients. IEEE Trans Inf Technol Biomed 2005;9:468-74.

300. Chen L, Reisner AT, McKenna TM, et al. Diagnosis of hemorrhage in a prehospital trauma population using linear and nonlinear multiparameter analysis of vital signs. Conf Proc IEEE Eng Med Biol Soc 2007;2007:3748-51.

301. Chen L, McKenna TM, Reisner AT, et al. Decision tool for the early diagnosis of trauma patient hypovolemia. J Biomed Inform 2008;41:469-78.

302. Convertino VA, Grudic G, Mulligan J, Moulton S. Estimation of individual-specific progression to impending cardiovascular instability using arterial waveforms. J Appl Physiol 2013;115:1196-1202.

303. Convertino VA, Howard JT, Hinojosa-Laborde C, et al. Individual-Specific, Beat-to-beat Trending of Significant Human Blood Loss: The Compensatory Reserve. Shock 2015;44 Suppl 1:27-32.

304. Convertino VA, Schiller AM. Measuring the compensatory reserve to identify shock. J Trauma Acute Care Surg 2017;82:S57-S65.

305. Sjogren AR, Leo MM, Feldman J, Gwin JT. Image Segmentation and Machine Learning for Detection of Abdominal Free Fluid in Focused Assessment With Sonography for Trauma Examinations: A Pilot Study. J Ultrasound Med 2016;35:2501-2509.

306. Najarian K, Hakimzadeh R, Ward K, et al. Combining predictive capabilities of transcranial doppler with electrocardiogram to predict hemorrhagic shock. Conf Proc IEEE Eng Med Biol Soc 2009;2009:2621-4.

307. Christie SA, Conroy AS, Callcut RA, et al. Dynamic multi-outcome prediction after injury: Applying adaptive machine learning for precision medicine in trauma. PLoS One 2019;14:e0213836.

308. Nadler R, Convertino VA, Gendler S, et al. The value of noninvasive measurement of the compensatory reserve index in monitoring and triage of patients experiencing minimal blood loss. Shock 2014;42:93-8.

309. The L. Artificial intelligence in health care: within touching distance. The Lancet 2017;390.

310. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44-56.

311. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018;1:18.

312. Castelvecchi D. Can we open the black box of AI? Nature 2016;538:20-23.

313. Ross C, Swetlitz I. IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. Stat News. Available at: https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/. Accessed 27 July 2019.

314. McCartney M. Margaret McCartney: Innovation without sufficient evidence is a disservice to all. BMJ 2017;358:j3980.

315. Oliver D. David Oliver: Lessons from the Babylon Health saga. BMJ 2019;365:l2387.

316. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. NPJ Digit Med 2018;1:40.

317. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. Lancet 2019;393:1577-1579.

318. Steiner DF, MacDonald R, Liu Y, et al. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. Am J Surg Pathol 2018;42:1636-1646.

319. McGrayne SB. The theory that would not die: how Bayes' rule cracked the enigma code, hunted down Russian submarines, & emerged triumphant from two centuries of controversy: Yale University Press; 2011.

320. Fenton N, Neil M. Risk assessment and decision analysis with Bayesian networks: Crc Press; 2012.

321. Efron B. Bayesians, frequentists, and scientists. Journal of the American Statistical Association 2005;100:1-5.

322. Bayes T, Price R, Canton J. An essay towards solving a problem in the doctrine of chances: C. Davis, Printer to the Royal Society of London; 1763.

323. Ltd A. AgenaRisk: Bayesian network and simulation software for risk analysis and decision support. Available at: www.AgenaRisk.com. Accessed 01/08/2017.

324. Piccininni M, Konigorski S, Rohmann JL, Kurth T. Directed acyclic graphs and causal thinking in clinical risk prediction modeling. BMC Med Res Methodol 2020;20:179.

325. Abdulkareem SA, Mustafa YT, Augustijn E-W, Filatova T. Bayesian networks for spatial learning: a workflow on using limited survey data for intelligent learning in spatial agent-based models. GeoInformatica 2019;23:243-268.

326. Pearl J. Causal inference in statistics: An overview. Statistics Surveys 2009;3:96-146.

327. Kyrimi E. Bayesian Networks for Clinical Decision Making: Support, Assurance, Trust: QMUL; 2019.

328. Uusitalo L. Advantages and challenges of Bayesian networks in environmental modelling. Ecol Modell 2007;203:312-318.

329. Lucas PJ, van der Gaag LC, Abu-Hanna A. Bayesian networks in biomedicine and health-care. Artif Intell Med 2004;30:201-14.

330. Yet B, Bastani K, Raharjo H, et al. Decision support system for Warfarin therapy management using Bayesian networks. Decis Support Syst 2013;55:488-498.

331. Lucas PJ, de Bruijn NC, Schurink K, Hoepelman A. A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. Artif Intell Med 2000;19:251-279.

332. Zou M, Conzen SD. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. Bioinformatics 2005;21:71-9.

333. Burnside ES, Rubin DL, Fine JP, et al. Bayesian network to predict breast cancer risk of mammographic microcalcifications and reduce number of benign biopsy results: initial experience. Radiology 2006;240:666-73.

334. McLachlan S, Dube K, Hitman GA, et al. Bayesian networks in healthcare: Distribution by medical condition. Artif Intell Med 2020;107:101912.

335. Yet B, Perkins Z, Fenton N, et al. Not just data: a method for improving prediction with knowledge. J Biomed Inform 2014;48:28-37.

336. Yet B. Bayesian Networks for Evidence Based Clinical Decision Support: QMUL; 2013.

337. Perkins Z. Bayesian Network Decision-Support for Severe Lower Limb Trauma: Queen Mary University of London; 2015.

338. Perkins ZB, Yet B, Marsden M, et al. Early Identification of Trauma-induced Coagulopathy: Development and Validation of a Multivariable Risk Prediction Model. Ann Surg 2020.

339. Brohi K, S. E. Traumatic coagulopathy and massive transfusion: improving outcomes and saving blood. Southampton (UK):: NIHR Journals Library; 2017 Nov. (Programme Grants for Applied Research, No. 5.19.); 2017 Nov 2017.

340. Yet B, Perkins ZB, Tai NRM, Marsh DWR. Clinical evidence framework for Bayesian networks. Knowl Inf Syst 2016;50:117-143.

341. Mossadegh. S. Application and Development of Bayesian Networks for Predictive Modelling of Coagulopathy and Mortality in Trauma Patients: QMUL; 2019.

342. Brohi K, Cohen MJ, Davenport RA. Acute coagulopathy of trauma: mechanism, identification and effect. Curr Opin Crit Care 2007;13:680-5.

343. Eastridge BJ, Malone D, Holcomb JB. Early predictors of transfusion and mortality after injury: a review of the data-based literature. J Trauma 2006;60:S20-5.

344. Maegele M, Lefering R, Yucel N, et al. Early coagulopathy in multiple injury: an analysis from the German Trauma Registry on 8724 patients. Injury 2007;38:298-304.

345. Collaborators C-. The importance of early treatment with tranexamic acid in bleeding trauma patients: an exploratory analysis of the CRASH-2 randomised controlled trial. The Lancet 2011;377:1096-1101. e2.

346. Gruen RL, Brohi K, Schreiber M, et al. Haemorrhage control in severely injured patients. Lancet 2012;380:1099-108.

347. Sorensen B, Fries D. Emerging treatment strategies for trauma-induced coagulopathy. Br J Surg 2012;99 Suppl 1:40-50.

348. Meyer DE, Vincent LE, Fox EE, et al. Every minute counts: Time to delivery of initial massive transfusion cooler and its impact on mortality. Journal of Trauma and Acute Care Surgery 2017;83:19-24.

349. Scott R, Jr., Crosby WH. Changes in the coagulation mechanism following wounding and resuscitation with stored blood; a study of battle casualties in Korea. Blood 1954;9:609-21.

350. Simmons RL, Collins JA, Heisterkamp CA, et al. Coagulation disorders in combat casualties. I. Acute changes after wounding. II. Effects of massive transfusion. 3. Post-resuscitative changes. Ann Surg 1969;169:455-82.

351. Ritchie J, Lewis J, Nicholls CM, Ormston R. Qualitative research practice: A guide for social science students and researchers. London: Sage; 2013.

352. Ormston R, Spencer L, Barnard M, Snape D. The foundations of qualitative research; 2014.

353. O'Brien BC, Harris IB, Beckman TJ, et al. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014;89:1245-51.

354. Seale C. The quality of qualitative research. London ; Thousand Oaks, Calif.: Sage Publications; 1999.

355. Caelli K, Ray L, Mill J. 'Clear as Mud': Toward Greater Clarity in Generic Qualitative Research. International Journal of Qualitative Methods 2003;2:1-13.

356. Cooper S, Endacott R. Generic qualitative research: a design for qualitative research in emergency care? Emerg Med J 2007;24:816-9.

357. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010;33:77-84.

358. Sullivan-Bolyai S, Bova C, Harper D. Developing and refining interventions in persons with health disparities: the use of qualitative description. Nurs Outlook 2005;53:127-33.

359. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23:334-40.

360. Bradshaw C, Atkinson S, Doody O. Employing a Qualitative Description Approach in Health Care Research. Glob Qual Nurs Res 2017;4:2333393617742282.

361. Percy W, Kostere K, Kostere S. Generic Qualitative Research in Psychology. The Qualitative Report 2015;20:76-85.

362. Britten N. Qualitative interviews in medical research. BMJ 1995;311:251-3.

363. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006;3:77-101.

364. Taylor S, Bogdan R, DeVault M. Introduction to qualitative research methods. 4 ed: Wiley; 2016.

365. Vaismoradi M, Jones J, Turunen H, Snelgrove S. Theme development in qualitative content analysis and thematic analysis. Journal of Nursing Education and Practice 2016;6.

366. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

367. Clarke V, Braun VJTp. Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. 2013;26:120-123.

368. Koffka K. Principles of Gestalt psychology: Routledge; 1935.

369. Klein JG. Five pitfalls in decisions about diagnosis and prescribing. BMJ 2005;330:781-3.

370. Steindl C, Jonas E, Sittenthaler S, et al. Understanding Psychological Reactance: New Developments and Findings. Z Psychol 2015;223:205-214.

371. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956;63:81-97.

372. Sacks GD, Dawes AJ, Ettner SL, et al. Surgeon Perception of Risk and Benefit in the Decision to Operate. Ann Surg 2016;264:896-903.

373. Simpson PA. Naturalistic Decision Making in Aviation Environments. Australia: Defence Science & Technology Organisation; 2001 January 2001.

374. Carvalho PVR, dos Santos IL, Vidal MCR. Nuclear power plant shift supervisor's decision making during microincidents. Int J Ind Ergon 2005;35:619-644.

375. Greitzer F, Podmore L. Naturalistic Decision Making in Power Grid Operations: Implications for Dispatcher Training and Usability Testing: Pacific Northwest National Laboratory; 2008 November 2008. Report No.: DE-AC05-76RL01830.

376. Pierluissi E, Fischer MA, Campbell AR, Landefeld CS. Discussion of medical errors in morbidity and mortality conferences. JAMA 2003;290:2838-42.

377. Sinitsky DM, Gowda SB, Dawas K, Fernando BS. Morbidity and mortality meetings to improve patient safety: a survey of 109 consultant surgeons in London, United Kingdom. Patient Saf Surg 2019;13:27.

378. Cialdini RB, Cialdini RB. Influence: The psychology of persuasion. 1993.

379. Elwyn G, Edwards A, Eccles M, Rovner D. Decision analysis in patient care. Lancet 2001;358:571-4.

380. Croskerry P, Nimmo GR. Better clinical decision making and reducing diagnostic error. J R Coll Physicians Edinb 2011;41:155-62.

381. Fraser J. The morbidity and mortality meeting: time for a different approach? Arch Dis Child 2016;101:4-8.

382. Rossi F. Building Trust in Artificial Intelligence. Journal of International Affairs 2018;72:127-134.

383. Lincoln YS, Guba EG. Naturalistic inquiry. Beverly Hills, Calif.: Sage Publications; 1985.

384. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52:1893-1907.

385. Klein GA, Calderwood R, MacGregor D. Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics 1989;19:462-472.

386. Committee of Medical Aspects of Automotive Safety. Rating the severity of tissue damage. I. The abbreviated scale. JAMA 1971;215:277-80.

387. Stanworth SJ, Davenport R, Curry N, et al. Mortality from trauma haemorrhage and opportunities for improvement in transfusion practice. Br J Surg 2016;103:357-65.

388. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32-5.

389. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.

390. Meyer DE, Cotton BA, Fox EE, et al. A comparison of resuscitation intensity and critical administration threshold in predicting early mortality among bleeding patients: A multicenter validation in 680 major transfusion patients. J Trauma Acute Care Surg 2018;85:691-696.

391. Savage SA, Zarzaur BL, Croce MA, Fabian TC. Redefining massive transfusion when every second counts. J Trauma Acute Care Surg 2013;74:396-400; discussion 400-2.

392. Rahbar E, Fox EE, del Junco DJ, et al. Early resuscitation intensity as a surrogate for bleeding severity and early mortality in the PROMMTT study. J Trauma Acute Care Surg 2013;75:S16-23.

393. Knaus WA, Harrell FE, Jr., Lynn J, et al. The SUPPORT prognostic model. Objective estimates of survival for seriously ill hospitalized adults. Study to understand prognoses and preferences for outcomes and risks of treatments. Ann Intern Med 1995;122:191-203.

394. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). Ann Intern Med 2015;162:735-6.

395. Frith D, Goslings JC, Gaarder C, et al. Definition and drivers of acute traumatic coagulopathy: clinical and experimental investigations. J Thromb Haemost 2010;8:1919-25.

396. Lauritzen SL. The EM algorithm for graphical association models with missing data. Computational Statistics & Data Analysis 1995;19:191-201.

397. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. Stat Med 2014;33:517-35.

398. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.

399. Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. Biom J 2008;50:457-79.

400. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565-74.

401. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ 2016;352:i6.

402. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. J Clin Oncol 2016;34:2534-40.

403. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.

404. Wickham H. ggplot2: elegant graphics for data analysis: Springer; 2016.

405. Harrell Jr FE. Package 'rms'. 2019.

406. Ter Avest E, Griggs J, Wijesuriya J, et al. Determinants of prehospital lactate in trauma patients: a retrospective cohort study. BMC Emerg Med 2020;20:18.

407. Silver N. The signal and the noise : why so many predictions fail--but some don't. New York: Penguin Press; 2012.

408. Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Med Res Methodol 2014;14:40.

409. Wallace E, Smith SM, Perera-Salazar R, et al. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). BMC Med Inform Decis Mak 2011;11:62.

410. Liu H, Hussain F, Tan CL, Dash M. Discretization: An enabling technique. Data Min Knowl Disc 2002;6:393-423.

411. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377-81.

412. Baksaas-Aasen K, Van Dieren S, Balvers K, et al. Data-driven Development of ROTEM and TEG Algorithms for the Management of Trauma Hemorrhage: A Prospective Observational Multicenter Study. Ann Surg 2019;270:1178-1185.

413. Obuchowski NA, Lieber ML, Wians FH, Jr. ROC curves in clinical chemistry: uses, misuses, and possible solutions. Clin Chem 2004;50:1118-25.

414. Obuchowski N. ROC Analysis. American Roentgen Ray Society; 2005.

415. Goksuluk D, Korkmaz S, Zararsiz G, Karaagaoglu A. easyROC: An InteractiveWeb-tool for ROC Curve Analysis Using R Language Environment. The R Journal 2016;8:213-230.

416. Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas 2016;20:37-46.

417. Cole E, Weaver A, Gall L, et al. A Decade of Damage Control Resuscitation: New Transfusion Practice, New Survivors, New Directions. Ann Surg 2021;273:1215-1220.

418. Surowiecki J. The wisdom of crowds : why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations: New York ; London : Doubleday : 2004.

419. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020;368:m689.

420. Association for the Advancement of Automotive Medicine e. The Abbreviated Injury Scale 2005. Des Plaines, IL: Association for the Advancement of Automotive Medicine; 2005.

421. Hasler RM, Kehl C, Exadaktylos AK, et al. Accuracy of prehospital diagnosis and triage of a Swiss helicopter emergency medical service. J Trauma Acute Care Surg 2012;73:709-15.

422. Augustin P, Guivarch E, Pellenc Q, et al. Performance of pre-hospital evaluations in ruling out invasive chest stab wounds. Scand J Trauma Resusc Emerg Med 2020;28:33.

423. ter Avest E, Taylor S, Wilson M, Lyon RL. Prehospital clinical signs are a poor predictor of raised intracranial pressure following traumatic brain injury. Emerg Med J 2020.

424. Buduhan G, McRitchie DI. Missed injuries in patients with multiple trauma. J Trauma 2000;49:600-5.

425. Giannakopoulos GF, Saltzherr TP, Beenen LF, et al. Missed injuries during the initial assessment in a cohort of 1124 level-1 trauma patients. Injury 2012;43:1517-21.

426. Biffl WL, Harrington DT, Cioffi WG. Implementation of a tertiary trauma survey decreases missed injuries. J Trauma 2003;54:38-43; discussion 43-4.

427. Yong E, Vasireddy A, Pavitt A, et al. Pre-hospital pelvic girdle injury: Improving diagnostic accuracy in a physician-led trauma service. Injury 2016;47:383-8.

428. Mazahir S, Pardhan A, Rao S. Office hours vs after-hours. Do presentation times affect the rate of missed injuries in trauma patients? Injury 2015;46:610-5.

429. Kyrimi E, Neves MR, McLachlan S, et al. Medical idioms for clinical Bayesian network development. J Biomed Inform 2020;108:103495.

430. Rothman KJ. Causes. Am J Epidemiol 1976;104:587-92.

431. Russo F, Williamson J. Epistemic causality and evidence-based medicine. Hist Philos Life Sci 2011;33:563-81.

432. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health 2005;95 Suppl 1:S144-50.

433. Hill AB. The Environment and Disease: Association or Causation? Proc R Soc Med 1965;58:295-300.

434. Howick J, Glasziou P, Aronson JK. The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? J R Soc Med 2009;102:186-94.

435. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. British Medical Journal 2009;339:b2535.

436. Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001;323:224-8.

437. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

438. Hooijmans CR, Rovers MM, de Vries RBM, et al. SYRCLE's risk of bias tool for animal studies. Bmc Medical Research Methodology 2014;14:43.

439. ACS. ATLS, Advanced Trauma Life Support for doctors. 9th ed. Chicago: American College of Surgeons; 2012.

440. Cohen MJ, Kutcher M, Redick B, et al. Clinical and mechanistic drivers of acute traumatic coagulopathy. J Trauma Acute Care Surg 2013;75:S40-7.

441. Talving P, Benfield R, Hadjizacharia P, et al. Coagulopathy in severe traumatic brain injury: a prospective study. J Trauma 2009;66:55-61; discussion 61-2.

442. Wafaisade A, Wutzler S, Lefering R, et al. Drivers of acute coagulopathy after severe trauma: a multivariate analysis of 1987 patients. Emerg Med J 2010;27:934-9.

443. Cotton BA, Sixta SL, Wade CE, et al. Mechanistic Determinants of the Acute Coagulopathy of Trauma (ACoT) in Patients Undergoing Emergency Surgery. Blood 2011;118:1427-1428.

444. Kashuk J. Primary fibrinolysis is intergral to the pathogenesis of the acute coagulopathy of trauma. Shock 2012;38:6.

445. Raza I, Davenport R, Rourke C, et al. The incidence and magnitude of fibrinolytic activation in trauma patients. Journal of Thrombosis and Haemostasis 2013;11:307-314.

446. van Zyl N, Milford EM, Diab S, et al. Activation of the protein C pathway and endothelial glycocalyx shedding is associated with coagulopathy in an ovine model of trauma and hemorrhage. J Trauma Acute Care Surg 2016;81:674-84.

447. Kapsch DN, Metzler M, Harrington M, et al. Fibrinolytic response to trauma. Surgery 1984;95:473-8.

448. Davis JW, Parks SN, Kaups KL, et al. Admission base deficit predicts transfusion requirements and risk of complications. The Journal of trauma 1996;41:769-74.

449. Brohi K, Cohen MJ, Ganter MT, et al. Acute traumatic coagulopathy: Initiated by hypoperfusion: Modulated through the protein C pathway? Annals of Surgery 2007;245:812-818.

450. Cohen MJ, Brohi K, Ganter MT, et al. Early coagulopathy after traumatic brain injury: the role of hypoperfusion and the protein C pathway. J Trauma 2007;63:1254-61; discussion 1261-2.

451. Brohi K, Cohen MJ, Ganter MT, et al. Acute coagulopathy of trauma: hypoperfusion induces systemic anticoagulation and hyperfibrinolysis. J Trauma 2008;64:1211-7; discussion 1217.

452. Cap AP, Spinella PC. Severity of head injury is associated with increased risk of coagulopathy in combat casualties. J Trauma 2011;71:S78-81.

453. Simmons JW, White CE, Ritchie JD, et al. Mechanism of injury affects acute coagulopathy of trauma in combat casualties. J Trauma 2011;71:S74-7.

454. Sixta SL, Hatch QM, Matijevic N, et al. Mechanistic determinates of the acute coagulopathy of trauma (ACoT) in patients requiring emergency surgery. International Journal of Burns & Trauma 2012;2:158-66.

455. Cheddie S, Muckart DJ, Hardcastle TC. Base deficit as an early marker of coagulopathy in trauma. S Afr J Surg 2013;51:88-90.

456. Xu SX, Wang L, Zhou GJ, et al. Risk factors and clinical significance of trauma-induced coagulopathy in ICU patients with severe trauma. Eur J Emerg Med 2013;20:286-90.

457. Deras P, Villiet M, Manzanera J, et al. Early coagulopathy at hospital admission predicts initial or delayed fibrinogen deficit in severe trauma patients. J Trauma Acute Care Surg 2014;77:433-40.

458. Lee TH, Hampton DA, Diggs BS, et al. Traumatic brain injury is not associated with coagulopathy out of proportion to injury in other body regions. J Trauma Acute Care Surg 2014;77:67-72; discussion 72.

459. White NJ, Contaifer D, Jr., Martin EJ, et al. Early hemostatic responses to trauma identified with hierarchical clustering analysis. J Thromb Haemost 2015;13:978-88.

460. Martini WZ, Chinkes DL, Pusateri AE, et al. Acute changes in fibrinogen metabolism and coagulation after hemorrhage in pigs. Am J Physiol Endocrinol Metab 2005;289:E930-4.

461. White NJ, Martin EJ, Brophy DF, Ward KR. Coagulopathy and traumatic shock: Characterizing hemostatic function during the critical period prior to fluid resuscitation. Resuscitation 2010;81:111-116.

462. Harr JN, Moore EE, Wohlauer MV, et al. The acute coagulopathy of trauma is due to impaired initial thrombin generation but not clot formation or clot strength. J Surg Res 2011;170:319-24.

463. Burruss S, Andakyan A, Romanov S, et al. Effect of protein C gene mutation on coagulation and inflammation in hemorrhagic shock. J Surg Res 2012;175:18-23.

464. Darlington DN, Kremenevskiy I, Pusateri AE, et al. Effects of In vitro hemodilution, hypothermia and rFVIIa addition on coagulation in human blood. International journal of burns and trauma 2012;2:42-50.

465. Frith D, Cohen MJ, Brohi K. Animal models of trauma-induced coagulopathy. Thromb Res 2012;129:551-6.

466. Letson HL, Pecheniuk NM, Mhango LP, Dobson GP. Reversal of acute coagulopathy during hypotensive resuscitation using small-volume 7.5% NaCl adenocaine and Mg2+ in the rat model of severe hemorrhagic shock. Crit Care Med 2012;40:2417-2422.

467. Park KH, Lee KH, Kim H. Effect of hypothermia on coagulatory function and survival in Sprague-Dawley rats exposed to uncontrolled haemorrhagic shock. Injury 2013;44:91-96.

468. Darlington DN, Craig T, Gonzales MD, et al. Acute coagulopathy of trauma in the rat. Shock 2013;39:440-6.

469. Frith D, Baptista ML, Cohen M, et al. Experimental acute traumatic coagulopathy is due to systemic activation of protein c during hypoperfusion. J Thromb Haemost 2011;9:481.

470. Jansen JO, Scarpelini S, Pinto R, et al. Hypoperfusion in Severely Injured Trauma Patients is Associated With Reduced Coagulation Factor Activity. Journal of Trauma-Injury Infection and Critical Care 2011;71:S435-S440.

471. Johansson PI, Sorensen AM, Perner A, et al. Disseminated intravascular coagulation or acute coagulopathy of trauma shock early after trauma? An observational study. Critical Care (London, England) 2011a;15.

472. Campbell JE, Meledeo MA, Cap AP. Comparative response of platelet fV and plasma fV to activated protein C and relevance to a model of acute traumatic coagulopathy. PLoS One 2014;9:e99181.

473. Hrafnkelsdottir T, Erlinge D, Jern S. Extracellular nucleotides ATP and UTP induce a marked acute release of tissue-type plasminogen activator in vivo in man. Thromb Haemost 2001;85:875-81.

474. Chapman MP, Moore EE, Moore HB, et al. Overwhelming tPA release, not PAI-1 degradation, is responsible for hyperfibrinolysis in severely injured trauma patients. J Trauma Acute Care Surg 2016;80:16-23; discussion 23-5.

475. Senzolo M, Coppell J, Cholongitas E, et al. The effects of glycosaminoglycans on coagulation: a thromboelastographic study. Blood coagulation & fibrinolysis : an international journal in haemostasis and thrombosis 2007;18:227-36.

476. Cardenas JC, Matijevic N, Baer LA, et al. Elevated tissue plasminogen activator and reduced plasminogen activator inhibitor promote hyperfibrinolysis in trauma patients. Shock 2014;41:514-21.

477. Letson HL, Dobson GP. Differential contributions of platelets and fibrinogen to early coagulopathy in a rat model of hemorrhagic shock. Thromb Res 2016;141:58-65.

478. Matijevic N, Wang YW, Wade CE, et al. Cellular microparticle and thrombogram phenotypes in the Prospective Observational Multicenter Major Trauma Transfusion (PROMMTT) study: correlation with coagulopathy. Thromb Res 2014;134:652-8.

479. Sillesen M, Johansson PI, Rasmussen LS, et al. Platelet activation and dysfunction in a large-animal model of traumatic brain injury and hemorrhage. J Trauma Acute Care Surg 2013;74:1252-9.

480. Li R, Elmongy H, Sims C, Diamond SL. Ex vivo recapitulation of trauma-induced coagulopathy and preliminary assessment of trauma patient platelet function under flow using microfluidic technology. Journal of Trauma and Acute Care Surgery 2016;80:440-449.

481. Rezende-Neto JB, Alves RL, Carvalho M, Jr., et al. Vagus nerve stimulation improves coagulopathy in hemorrhagic shock: a thromboelastometric animal model study. Journal of trauma management & outcomes 2014;8:15.

482. Walsh TS, Stanworth SJ, Prescott RJ, et al. Prevalence, management, and outcomes of critically ill patients with prothrombin time prolongation in United Kingdom intensive care units*. Critical Care Medicine 2010;38:1939-1946.

483. Sherren P, Hussey J, Martin R, et al. Acute burn induced coagulopathy. Burns 2013;39:1157-1161.

484. Adrie C, Monchi M, Laurent I, et al. Coagulopathy after successful cardiopulmonary resuscitation following cardiac arrest: implication of the protein C anticoagulant pathway. J Am Coll Cardiol 2005;46:21-8.

485. White NJ, Leong BS, Brueckner J, et al. Coagulopathy during cardiac arrest and resuscitation in a swine model of electrically induced ventricular fibrillation. Resuscitation 2011;82:925-31.

486. Lee SH, Kim K, Kim J, et al. Factors Associated with Early Coagulopathy in Witnessed Sudden Cardiac Arrest. Critical Care Medicine 2012;40:U159-U159.

487. Schochl H, Cadamuro J, Seidl S, et al. Hyperfibrinolysis is common in out-of-hospital cardiac arrest: results from a prospective observational thromboelastometry study. Resuscitation 2013;84:454-9.

488. Banner DW, D'Arcy A, Chene C, et al. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Nature 1996;380:41-6.

489. Hoffman M, Monroe DM, 3rd. A cell-based model of hemostasis. Thromb Haemost 2001;85:958-65.

490. Niles SE, McLaughlin DF, Perkins JG, et al. Increased mortality associated with the early coagulopathy of trauma in combat casualties. J Trauma 2008;64:1459-63; discussion 1463-5.

491. Nardai G, Gombkoto A, Basta M, et al. Associations between complement system activation and acute traumatic coagulopathy. European Journal of Anaesthesiology 2009;26.

492. Shaz BH, Winkler AM, James AB, et al. Pathophysiology of early trauma-induced coagulopathy: emerging evidence for hemodilution and coagulation factor depletion. J Trauma 2011;70:1401-7.

493. Floccard B, Rugeri L, Faure A, et al. Early coagulopathy in trauma patients: an on-scene and hospital admission study. Injury 2012;43:26-32.

494. Genet GF, Johansson PI, Meyer MA, et al. Trauma-induced coagulopathy: standard coagulation tests, biomarkers of coagulopathy, and endothelial damage in patients with traumatic brain injury. J Neurotrauma 2013;30:301-6.

495. Lustenberger T, Talving P, Kobayashi L, et al. Early coagulopathy after isolated severe traumatic brain injury: relationship with hypoperfusion challenged. J Trauma 2010;69:1410-4.

496. Cordts MR, Parreira JG, Perlingeiro JA, et al. Pelvic fractures as a marker of injury severity in trauma patients. Revista do Colegio Brasileiro de Cirurgioes 2011;38:310-6.

497. Filho RMC, Parreira JG, Perlingeiro JAG, et al. Pelvic fractures as a marker of injury severity in trauma patients [English;Portuguese] Fratura de pelve: Um marcador de gravidade em trauma. Revista do Colegio Brasileiro de Cirurgioes 2011;38:310-316.

498. Rotondo MF, Schwab CW, McGonigal MD, et al. 'Damage control': an approach for improved survival in exsanguinating penetrating abdominal injury. Journal of Trauma - Injury, Infection and Critical Care 1993;35:375-383.

499. Garrison JR, Richardson JD, Hilakos AS, et al. Predicting the need to pack early for severe intra-abdominal hemorrhage. J Trauma 1996;40:923-7; discussion 927-9.

500. Ferrara A, MacArthur JD, Wright HK, et al. Hypothermia and acidosis worsen coagulopathy in the patient requiring massive transfusion. The American Journal of Surgery 1990;160:515-518.

501. Aucar JA, Norman P, Whitten E, et al. Intraoperative detection of traumatic coagulopathy using the activated coagulation time. Shock 2003;19:404-7.

502. Engels PT, Rezende-Neto JB, Al Mahroos M, et al. The natural history of trauma-related coagulopathy: implications for treatment. J Trauma 2011;71:S448-55.

503. Neal MD, Brown JB, Moore EE, et al. Prehospital use of nonsteroidal anti-inflammatory drugs (NSAIDs) is associated with a reduced incidence of trauma-induced coagulopathy. Ann Surg 2014;260:378-82.

504. Darlington DN, Kheirabadi BS, Delgado AV, et al. Coagulation changes to systemic acidosis and bicarbonate correction in swine. J Trauma 2011;71:1271-7.

505. Dubick M, Scherer M, Darlington D, Kheriabadi B. Effects of Acidosis and Ph Correction on Thrombin Generation in a Swine Hemorrhage Model. Critical Care Medicine 2009;37:A125-A125.

506. Martini WZ, Dubick MA, Pusateri AE, et al. Does bicarbonate correct coagulation function impaired by acidosis in swine? Journal of Trauma and Acute Care Surgery 2006a;61:99-106.

507. Engstrom M, Schott U, Romner B, Reinstrup P. Acidosis impairs the coagulation: A thromboelastographic study. J Trauma 2006;61:624-8.

508. Dirkmann D, Radu-Berlemann J, Gorlinger K, Peters J. Recombinant tissue-type plasminogen activator-evoked hyperfibrinolysis is enhanced by acidosis and inhibited by hypothermia but still can be blocked by tranexamic acid. J Trauma Acute Care Surg 2013;74:482-8.

509. Engström M, Schött U, Romner B, Reinstrup P. Acidosis impairs the coagulation: A thromboelastographic study. Journal of Trauma - Injury, Infection and Critical Care 2006;61:624-628.

510. Martini WZ, Dubick MA, Wade CE, Holcomb JB. Evaluation of tris-hydroxymethylaminomethane on reversing coagulation abnormalities caused by acidosis in pigs. Critical Care Medicine 2007;35:1568-74.

511. Martini WZ, Holcomb JB. Acidosis and coagulopathy: the differential effects on fibrinogen synthesis and breakdown in pigs. Annals of surgery 2007b;246:831-835.

512. Martini WZ, Pusateri AE, Uscilowicz JM, et al. Independent contributions of hypothermia and acidosis to coagulopathy in swine. Journal of Trauma - Injury, Infection and Critical Care 2005b;58:1002-1010.

513. Ireland S, Endacott R, Cameron P, et al. The incidence and significance of accidental hypothermia in major trauma--a prospective observational study. Resuscitation 2011;82:300-6.

514. Watts DD, Trask A, Soeken K, et al. Hypothermic coagulopathy in trauma: Effect of varying levels of hypothermia on enzyme speed, platelet function, and fibrinolytic activity. Journal of Trauma - Injury, Infection and Critical Care 1998;44:846-854.

515. Kashuk JL, Moore EE, Sawyer M, et al. Primary fibrinolysis is integral in the pathogenesis of the acute coagulopathy of trauma. Annals of surgery 2010;252:434-444.

516. Kettner SC, Sitzwohl C, Zimpfer M, et al. The effect of graded hypothermia (36 degrees C-32 degrees C) on hemostasis in anesthetized patients without surgical trauma. Anesth Analg 2003;96:1772-6, table of contents.

517. Whelihan MF, Kiankhooy A, Brummel-Ziedins KE. Thrombin generation and fibrin clot formation under hypothermic conditions: an in vitro evaluation of tissue factor initiated whole blood coagulation. J Crit Care 2014;29:24-30.

518. Wolberg AS, Meng ZH, Monroe DM, 3rd, Hoffman M. A systematic evaluation of the effect of temperature on coagulation enzyme activity and platelet function. Journal of Trauma - Injury, Infection and Critical Care 2004;56:1221-8.

519. Michelson AD, MacGregor H, Barnard MR, et al. Reversible inhibition of human platelet activation by hypothermia in vivo and in vitro. Thrombosis and haemostasis 1994;71:633-640.

520. Martini WZ, Cortez DS, Dubick MA, et al. Thrombelastography is better than PT, aPTT, and activated clotting time in detecting clinically relevant clotting abnormalities after hypothermia, hemorrhagic shock and resuscitation in pigs. Journal of Trauma - Injury, Infection and Critical Care 2008;65:535-543.

521. Martini WZ. The effects of hypothermia on fibrinogen metabolism and coagulation function in swine. Metabolism 2007a;56:214-21.

522. Martini WZ. The effects of hypothermia on fibrinogen metabolism and coagulation function in swine. Metabolism 2007;56:214-21.

523. Mitrophanov AY, Rosendaal FR, Reifman J. Computational analysis of the effects of reduced temperature on thrombin generation: the contributions of hypothermia to coagulopathy. Anesth Analg 2013;117:565-74.

524. Meng ZH, Wolberg AS, Monroe DM, 3rd, Hoffman M. The effect of temperature and pH on the activity of factor VIIa: implications for the efficacy of high-dose factor VIIa in hypothermic and acidotic patients. Journal of Trauma - Injury, Infection and Critical Care 2003;55:886-91.

525. Mohr J, Ruchholtz S, Hildebrand F, et al. Induced hypothermia does not impair coagulation system in a swine multiple trauma model. Journal of Trauma and Acute Care Surgery 2013;74:1014-1020.

526. Watts DD, Trask A, Soeken K, et al. Hypothermic coagulopathy in trauma: effect of varying levels of hypothermia on enzyme speed, platelet function, and fibrinolytic activity. J Trauma 1998;44:846-54.

527. Hubetamann B, Lefering R, Taeger G, et al. Influence of prehospital fluid resuscitation on patients with multiple injuries in hemorrhagic shock in patients from the DGU trauma registry. Journal of emergencies, trauma, and shock 2011;4:465-71.

528. Maegele M, Lefering R, Yucel N, et al. Early coagulopathy in multiple injury: an analysis from the German Trauma Registry on 8724 patients. Injury 2007;38:298-304.

529. Rourke C, Curry NS, Khan S, et al. Characterization of fibrinogen depletion and response to replacement therapy in bleeding trauma patients. Journal of Thrombosis and Haemostasis 2011;9.

530. Darlington D, Fedyk C, Scherer M, et al. Hemodilution Effects Coagulation of Human Blood Ex Vivo. Critical Care Medicine 2010;38:U37-U37.

531. Bolliger D, Szlam F, Levy JH, et al. Haemodilution-induced profibrinolytic state is mitigated by fresh-frozen plasma: implications for early haemostatic intervention in massive haemorrhage. Br J Anaesth 2010;104:318-25.

532. Ogweno GO, Gwer S. Haemodilution with Nacl solutions below 600mOsm/L Induce hypercoagulability while concentrations above 1200 mOsm/L progressively impair coagulation as evaluated with thrombelastography in vitro. J Thromb Haemost 2013;11:1028-1029.

533. Brazil EV, Coats TJ. Sonoclot coagulation analysis of in-vitro haemodilution with resuscitation solutions. J R Soc Med 2000;93:507-10.

534. Coats TJ, Brazil E, Heron M, MacCallum PK. Impairment of coagulation by commonly used resuscitation fluids in human volunteers. Emerg Med J 2006;23:846-9.

535. Caballo C, Escolar G, Diaz-Ricart M, et al. Impact of experimental haemodilution on platelet function, thrombin generation and clot firmness: effects of different coagulation factor concentrates. Blood Transfus 2013;11:391-9.

536. Dickneite G, Pragst I. Prothrombin complex concentrate vs fresh frozen plasma for reversal of dilutional coagulopathy in a porcine trauma model. Br J Anaesth 2009;102:345-54.

537. Dickneite G, Dorr B, Kaspereit F, Tanaka KA. Prothrombin complex concentrate versus recombinant factor VIIa for reversal of hemodilutional coagulopathy in a porcine trauma model. The Journal of trauma 2010;68:1151-7.

538. Grottke O, Braunschweig T, Henzler D, et al. Effects of different fibrinogen concentrations on blood loss and coagulation parameters in a pig model of coagulopathy with blunt liver injury. Crit Care 2010;14:R62.

539. Watts S, Nordmann G, Brohi K, et al. Evaluation of Prehospital Blood Products to Attenuate Acute Coagulopathy of Trauma in a Model of Severe Injury and Shock in Anesthetized Pigs. Shock 2015;44 Suppl 1:138-48.

540. Spronk HM, Braunschweig T, Rossaint R, et al. Recombinant Factor VIIa Reduces Bleeding after Blunt Liver Injury in a Pig Model of Dilutional Coagulopathy under Severe Hypothermia. PLoS One 2015;10:e0113979.

541. Hirshberg A, Dugas M, Banez EI, et al. Minimizing dilutional coagulopathy in exsanguinating hemorrhage: a computer simulation. J Trauma 2003;54:454-63.

542. Schols SE, Feijge MA, Lance MD, et al. Effects of plasma dilution on tissue-factor-induced thrombin generation and thromboelastography: partly compensating role of platelets. Transfusion 2008;48:2384-94.

543. Martini WZ, Chinkes DL, Sondeen J, Dubick MA. Effects of hemorrhage and lactated Ringer's resuscitation on coagulation and fibrinogen metabolism in swine. Shock 2006b;26:396-401.

544. Martini WZ, Dubick MA, Blackbourne LH. Comparisons of lactated Ringer's and Hextend resuscitation on hemodynamics and coagulation following femur injury and severe hemorrhage in pigs. J Trauma Acute Care Surg 2013;74:732-9; discussion 739-40.

545. Torres LN, Sondeen JL, Ji L, et al. Evaluation of resuscitation fluids on endothelial glycocalyx, venular blood flow, and coagulation function after hemorrhagic shock in rats. J Trauma Acute Care Surg 2013;75:759-66.

546. Fries D, Haas T, Klingler A, et al. Efficacy of fibrinogen and prothrombin complex concentrate used to reverse dilutional coagulopathy—a porcine model. BJA: British Journal of Anaesthesia 2006;97:460-467.

547. Fries D, Krismer A, Klingler A, et al. Effect of fibrinogen on reversal of dilutional coagulopathy: a porcine model. Br J Anaesth 2005;95:172-7.

548. Pragst I, Kaspereit F, Dorr B, Dickneite G. Prothrombin complex concentrate (Beriplex P/N) for control of bleeding after kidney trauma in a rabbit dilutional coagulopathy model. Thromb Res 2010;125:272-7.

549. Klemcke HG, Delgado A, Holcomb JB, et al. Effect of recombinant FVIIa in hypothermic, coagulopathic pigs with liver injuries. J Trauma 2005;59:155-61; discussion 161.

550. De Blainville HMD. Injection de matiere cerebrale dans des veines. Gazette Medicale de Paris 1834;2:524.

551. Harhangi BS, Kompanje EJ, Leebeek FW, Maas AI. Coagulation disorders after traumatic brain injury. Acta Neurochir (Wien) 2008;150:165-75; discussion 175.

552. Yuan Q, Sun YR, Wu X, et al. Coagulopathy in Traumatic Brain Injury and Its Correlation with Progressive Hemorrhagic Injury: A Systematic Review and Meta-Analysis. J Neurotrauma 2016;33:1279-91.

553. Goodnight SH, Kenoyer G, Rapaport SI, et al. Defibrination after brain-tissue destruction: A serious complication of head injury. N Engl J Med 1974;290:1043-7.

554. Auer L. Disturbances of the coagulatory system in patients with severe cerebral trauma. I. Acta Neurochir (Wien) 1978;43:51-9.

555. Bredbacka S, Edner G. Soluble Fibrin and D-Dimer as Detectors of Hypercoagulability in Patients with Isolated Brain Trauma. J Neurosurg Anesthesiol 1994;6:75-82.

556. Hulka F, Mullins RJ, Frank EH. Blunt brain injury activates the coagulation process. Arch Surg 1996;131:923-7; discussion 927-8.

557. Carrick MM, Tyroch AH, Youens CA, Handley T. Subsequent development of thrombocytopenia and coagulopathy in moderate and severe head injury: support for serial laboratory examination. The Journal of trauma 2005;58:725-9; discussion 729-30.

558. Zehtabchi S, Soghoian S, Liu Y, et al. The association of coagulopathy and traumatic brain injury in patients with isolated head injury. Resuscitation 2008;76:52-6.

559. Halpern CH, Reilly PM, Turtz AR, Stein SC. Traumatic coagulopathy: the effect of brain injury. J Neurotrauma 2008;25:997-1001.

560. Di Battista AP, Rizoli SB, Lejnieks B, et al. Sympathoadrenal Activation is Associated with Acute Traumatic Coagulopathy and Endotheliopathy in Isolated Brain Injury. Shock 2016;46:96-103.

561. Kearney TJ, Bentt L, Grode M, et al. Coagulopathy and catecholamines in severe head injury. The Journal of trauma 1992;32:608-11; discussion 611-2.

562. Kinoshita K. Traumatic brain injury: pathophysiology for neurocritical care. J Intensive Care 2016;4:29.

563. Lozance K, Dejanov I, Mircevski M. Role of coagulopathy in patients with head trauma. J Clin Neurosci 1998;5:394-8.

564. Keller MS, Fendya DG, Weber TR. Glasgow Coma Scale predicts coagulopathy in pediatric trauma patients. Semin Pediatr Surg 2001;10:12-6.

565. van der Sande JJ, Veltkamp JJ, Boekhout-Mussert RJ, Bouwhuis-Hoogerwerf ML. Head injury and coagulation disorders. J Neurosurg 1978;49:357-65.

566. Wafaisade A, Lefering R, Tjardes T, et al. Acute coagulopathy in isolated blunt traumatic brain injury. Neurocrit Care 2010;12:211-9.

567. Peiniger S, Nienaber U, Lefering R, et al. Glasgow Coma Scale as a predictor for hemocoagulative disorders after blunt pediatric traumatic brain injury. Pediatr Crit Care Med 2012;13:455-60.

568. Chhabra G, Sharma S, Subramanian A, et al. Coagulopathy as prognostic marker in acute traumatic brain injury. J Emerg Trauma Shock 2013;6:180-5.

569. Patel H, De La Cerda G, Healy M, Verma V. Association of coagulopathy with marshall ct scan classification in traumatic brain injury (TBI). Neurocrit Care 2011;15:S171.

570. Keimowitz RM, Annis BL. Disseminated intravascular coagulation associated with massive brain injury. J Neurosurg 1973;39:178-80.

571. Drake TA, Morrissey JH, Edgington TS. Selective cellular expression of tissue factor in human tissues. Implications for disorders of hemostasis and thrombosis. Am J Pathol 1989;134:1087-97.

572. Rapaport SI, Rao LV. The tissue factor pathway: how it has become a "prima ballerina". Thromb Haemost 1995;74:7-17.

573. Nekludov M, Antovic J, Bredbacka S, Blomback M. Coagulation abnormalities associated with severe isolated traumatic brain injury: cerebral arterio-venous differences in coagulation and inflammatory markers. J Neurotrauma 2007;24:174-80.

574. Cortiana M, Zagara G, Fava S, Seveso M. Coagulation abnormalities in patients with head injury. J Neurosurg Sci 1986;30:133-8.

575. Donahue DL, Beck J, Fritz B, et al. Early platelet dysfunction in a rodent model of blunt traumatic brain injury reflects the acute traumatic coagulopathy found in humans. J Neurotrauma 2014;31:404-10.

576. Castellino FJ, Chapman MP, Donahue DL, et al. Traumatic brain injury causes platelet adenosine diphosphate and arachidonic acid receptor inhibition independent of hemorrhagic shock in humans and rats. J Trauma Acute Care Surg 2014;76:1169-76.

577. Wu X, Du Z, Yu J, et al. Activity of factor VII in patients with isolated blunt traumatic brain injury: association with coagulopathy and progressive hemorrhagic injury. J Trauma Acute Care Surg 2014;76:114-20.

578. Morel N, Morel O, Petit L, et al. Generation of procoagulant microparticles in cerebrospinal fluid and peripheral blood after traumatic brain injury. J Trauma 2008;64:698-704.

579. Tian Y, Salsbery B, Wang M, et al. Brain-derived microparticles induce systemic coagulation in a murine model of traumatic brain injury. Blood 2015;125:2151-9.

580. Zhao Z, Wang M, Tian Y, et al. Cardiolipin-mediated procoagulant activity of mitochondria contributes to traumatic brain injury-associated coagulopathy in mice. Blood 2016;127:2763-72.

581. Gall LS, Vulliamy P, Gillespie S, et al. The S100A10 Pathway Mediates an Occult Hyperfibrinolytic Subtype in Trauma Patients. Ann Surg 2019;269:1184-1191.

582. Kaczynski J, Wilczynska M, Hilton J, Fligelstone L. Impact of crystalloids and colloids on coagulation cascade during trauma resuscitation-a literature review. Emergency Medicine and Health Care 2013;1:1.

583. Rossaint R, Bouillon B, Cerny V, et al. The European guideline on management of major bleeding and coagulopathy following trauma: fourth edition. Crit Care 2016;20:100.

584. Perel P, Roberts I, Ker K. Colloids versus crystalloids for fluid resuscitation in critically ill patients. Cochrane Database Syst Rev 2013:CD000567.

585. Zarychanski R, Abou-Setta AM, Turgeon AF, et al. Association of hydroxyethyl starch administration with mortality and acute kidney injury in critically ill patients requiring volume resuscitation: a systematic review and meta-analysis. JAMA 2013;309:678-88.

586. Coats TJ, Brazil E, Heron M. The effects of commonly used resuscitation fluids on whole blood coagulation. Emerg Med J 2006;23:546-9.

587. Kind SL, Spahn-Nett GH, Emmert MY, et al. Is dilutional coagulopathy induced by different colloids reversible by replacement of fibrinogen and factor XIII concentrates? Anesth Analg 2013;117:1063-71.

588. Sossdorf M, Marx S, Schaarschmidt B, et al. HES 130/0.4 impairs haemostasis and stimulates pro-inflammatory blood platelet function. Crit Care 2009;13:R208.

589. Evans PA, Garnett M, Boffard K, et al. Evaluation of the effect of colloid (Haemaccel) on the bleeding time in the trauma patient. J R Soc Med 1996;89:101P-4P.

590. Ekseth K, Abildgaard L, Vegfors M, et al. The in vitro effects of crystalloids and colloids on coagulation. Anaesthesia 2002;57:1102-8.

591. Morris BR, deLaforcade A, Lee J, et al. Effects of in vitro hemodilution with crystalloids, colloids, and plasma on canine whole blood coagulation as determined by kaolin-activated thromboelastography. J Vet Emerg Crit Care (San Antonio) 2016;26:58-63.

592. Brummel-Ziedins K, Whelihan MF, Ziedins EG, Mann KG. The resuscitative fluid you choose may potentiate bleeding. The Journal of trauma 2006;61:1350-8.

593. Roche AM, James MF, Bennett-Guerrero E, Mythen MG. A head-to-head comparison of the in vitro coagulation effects of saline-based and balanced electrolyte crystalloid and colloid intravenous fluids. Anesth Analg 2006;102:1274-9.

594. Mardel SN, Saunders FM, Allen H, et al. Reduced quality of clot formation with gelatin-based plasma substitutes. Br J Anaesth 1998;80:204-7.

595. Innerhofer P, Fries D, Margreiter J, et al. The effects of perioperatively administered colloids and crystalloids on primary platelet-mediated hemostasis and clot formation. Anesth Analg 2002;95:858-65, table of contents.

596. Delano MJ, Rizoli SB, Rhind SG, et al. Prehospital Resuscitation of Traumatic Hemorrhagic Shock with Hypertonic Solutions Worsens Hypocoagulation and Hyperfibrinolysis. Shock 2015;44:25-31.

597. Tan TS, Tan KH, Ng HP, Loh MW. The effects of hypertonic saline solution (7.5%) on coagulation and fibrinolysis: an in vitro assessment using thromboelastography. Anaesthesia 2002;57:644-8.

598. Coats TJ, Heron M. The effect of hypertonic saline dextran on whole blood coagulation. Resuscitation 2004;60:101-4.

599. Smith CA, Gosselin RC, Utter GH, et al. Does saline resuscitation affect mechanisms of coagulopathy in critically ill trauma patients? An exploratory analysis. Blood Coagul Fibrinolysis 2015;26:250-4.

600. Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning. Nat Commun 2020;11:3923.

601. Kyrimi E, Mossadegh S, Tai N, Marsh W. An incremental explanation of inference in Bayesian networks for increasing model trustworthiness and supporting clinical decision making. Artif Intell Med 2020;103:101812.

602. Select Committee on Artificial Intelligence. The National Artificial Intelligence Research and Development Strategic Plan: 2019 update: National Science and Technology Council (US); 2019.

603. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The lancet digital health 2019;1:e271-e297.

604. Wilkinson J, Arnold KF, Murray EJ, et al. Time to reality check the promises of machine learning-powered precision medicine. The Lancet Digital Health 2020.

605. Sendak M, Elish MC, Gao M, et al. "The human body is a black box". Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency; 2020. p. 99-109.

606. Wynants L, van Smeden M, McLernon DJ, et al. Three myths about risk thresholds for prediction models. BMC Med 2019;17:192.

607. Dewinter G, Staelens W, Veef E, et al. Simplified algorithm for the prevention of postoperative nausea and vomiting: a before-and-after study. Br J Anaesth 2018;120:156-163.

608. Snooks HA, Carter B, Dale J, et al. Support and Assessment for Fall Emergency Referrals (SAFER 1): cluster randomised trial of computerised clinical decision support for paramedics. PLoS One 2014;9:e106436.

609. Hagiwara MA, Sjoqvist BA, Lundberg L, et al. Decision support system in prehospital care: a randomized controlled simulation study. Am J Emerg Med 2013;31:145-53.

610. Vincent JL, Quintairos ESA, Couto L, Jr., Taccone FS. The value of blood lactate kinetics in critically ill patients: a systematic review. Crit Care 2016;20:257.

611. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010;43:159-72.

612. Neves MR, Marsh DWR. Modelling the Impact of AI for Clinical Decision Support. 2019; Cham: Springer International Publishing; 2019. p. 292-297.

613. Vellido A. The importance of interpretability and visualization in machine learning for applications in medicine and health care. Neural Computing and Applications 2019.

614. Miller T. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence 2019;267:1-38.

615. Kyrimi E, Dube K, Fenton N, et al. Bayesian networks in healthcare: What is preventing their adoption? Artif Intell Med 2021;116:102079.

616. Kyrimi E, McLachlan S, Dube K, et al. A comprehensive scoping review of Bayesian networks in healthcare: Past, present and future. Artif Intell Med 2021;117:102108.

617. DECIDE-AI Steering Group. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence. Nat Med 2021;27:186-187.

618. Rivera SC, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. BMJ 2020;370:m3210.

619. Brown JB, Cohen MJ, Minei JP, et al. Characterization of acute coagulopathy and sexual dimorphism after injury: females and coagulopathy just do not mix. J Trauma Acute Care Surg 2012;73:1395-400.

620. Brown JB, Cohen MJ, Minei JP, et al. Goal-directed resuscitation in the prehospital setting: a propensity-adjusted analysis. J Trauma Acute Care Surg 2013;74:1207-12; discussion 1212-4.

621. Charbit J, Lakhal K, Deras P, et al. Influence of surgical bleeding on the relationship between admission coagulopathy and risk of massive transfusion: lesson from 704 severe trauma patients. Vox Sang 2016;111:151-60.

622. Cohen MJ, Brohi K, Calfee CS, et al. Early release of high mobility group box nuclear protein 1 after severe trauma in humans: role of injury severity and tissue hypoperfusion. Crit Care 2009;13:R174.

623. Hayakawa M, Sawamura A, Gando S, et al. Disseminated intravascular coagulation at an early phase of trauma is associated with consumption coagulopathy and excessive fibrinolysis both by plasmin and neutrophil elastase. Surgery 2011;149:221-30.

624. Johansson PI, Sorensen AM, Perner A, et al. High sCD40L levels early after trauma are associated with enhanced shock, sympathoadrenal activation, tissue and endothelial damage, coagulopathy and mortality. J Thromb Haemost 2012;10:207-16.

625. Kutcher ME, Ferguson AR, Cohen MJ. A principal component analysis of coagulation after trauma. J Trauma Acute Care Surg 2013;74:1223-9; discussion 1229-30.

626. White NJ, Newton JC, Martin EJ, et al. Clot Formation is Associated With Fibrinogen and Platelet Forces in a Cohort of Severely-Injured Emergency Department Trauma Patients. Shock 2015a;44:39-44.

627. Chen J, Wu X, Keesee J, et al. Limited Resuscitation With Fresh or Stored Whole Blood Corrects Cardiovascular and Metabolic Function in a Rat Model of Polytrauma and Hemorrhage. Shock 2017;47:208-216.

628. Cho SD, Holcomb JB, Tieu BH, et al. Reproducibility of an animal model simulating complex combat-related injury in a multiple-institution format. Shock 2009;31:87-96.

629. Darlington DN, Gonzales MD, Craig T, et al. Trauma-Induced Coagulopathy Is Associated with a Complex Inflammatory Response in the Rat. Shock 2015;44 Suppl 1:129-37.

630. Doran CM, Doran CA, Woolley T, et al. Targeted resuscitation improves coagulation and outcome. J Trauma Acute Care Surg 2012;72:835-43.

631. Duan K, Yu W, Lin Z, et al. A time course study of acute traumatic coagulopathy prior to resuscitation: from hypercoagulation to hypocoagulation caused by hypoperfusion? Transfus Apher Sci 2014;50:399-406.

632. Fung YL, Tung JP, Foley SR, et al. Stored blood transfusion induces transient pulmonary arterial hypertension without impairing coagulation in an ovine model of nontraumatic haemorrhage. Vox Sang 2013;105:150-8.

633. Hagemo JS, Jorgensen JJ, Ostrowski SR, et al. Changes in fibrinogen availability and utilization in an animal model of traumatic coagulopathy. Scand J Trauma Resusc Emerg Med 2013;21:56.

634. Iwamoto S, Takasu A, Sakamoto T. Therapeutic mild hypothermia: effects on coagulopathy and survival in a rat hemorrhagic shock model. J Trauma 2010;68:669-75.

635. Martini WZ, Chinkes DL, Sondeen J, Dubick MA. Effects of hemorrhage and lactated Ringer's resuscitation on coagulation and fibrinogen metabolism in swine. Shock 2006;26:396-401.

636. Nishi K, Takasu A, Shinozaki H, et al. Hemodilution as a result of aggressive fluid resuscitation aggravates coagulopathy in a rat model of uncontrolled hemorrhagic shock. Journal of Trauma and Acute Care Surgery 2013;74:808-812.

637. Xu L, Yu WK, Lin ZL, et al. Chemical sympathectomy attenuates inflammation, glycocalyx shedding and coagulation disorders in rats with acute traumatic coagulopathy. Blood Coagul Fibrinolysis 2015;26:152-60.

638. Xu L, Yu WK, Lin ZL, et al. Impact of A-adrenoceptor blockade on systemic inflammation and coagulation disturbances in rats with acute traumatic coagulopathy. Med Sci Monit 2015;21:468-476.

639. Cohen MJ, Carles M, Brohi K, et al. Early release of soluble receptor for advanced glycation endproducts after severe trauma in humans. J Trauma 2010;68:1273-8.

640. Dunbar NM, Chandler WL. Thrombin generation in trauma patients. Transfusion (Paris) 2009;49:2652-60.

641. Johansson PI, Stensballe J, Rasmussen LS, Ostrowski SR. A high admission syndecan-1 level, a marker of endothelial glycocalyx degradation, is associated with inflammation, protein C depletion, fibrinolysis, and increased mortality in trauma patients. Ann Surg 2011;254:194-200.

642. Johansson PI, Stensballe J, Rasmussen LS, Ostrowski SR. High circulating adrenaline levels at admission predict increased mortality after trauma. J Trauma Acute Care Surg 2012;72:428-36.

643. Johansson PI, Windelov NA, Rasmussen LS, et al. Blood levels of histone-complexed DNA fragments are associated with coagulopathy, inflammation and endothelial damage early after trauma. J Emerg Trauma Shock 2013;6:171-5.

644. Khan S, Brohi K, Chana M, et al. Hemostatic resuscitation is neither hemostatic nor resuscitative in trauma hemorrhage. J Trauma Acute Care Surg 2014;76:561-7; discussion 567-8.

645. Lustenberger T, Relja B, Puttkammer B, et al. Activated thrombin-activatable fibrinolysis inhibitor (TAFIa) levels are decreased in patients with trauma-induced coagulopathy. Thromb Res 2013;131:e26-30.

646. MacLeod JB, Winkler AM, McCoy CC, et al. Early trauma induced coagulopathy (ETIC): prevalence across the injury spectrum. Injury 2014;45:910-5.

647. May AK, Young JS, Butler K, et al. Coagulopathy in severe closed head injury: is empiric therapy warranted? Am Surg 1997;63:233-6; discussion 236-7.

648. Ostrowski SR, Sorensen AM, Larsen CF, Johansson PI. Thrombelastography and biomarker profiles in acute coagulopathy of trauma: a prospective study. Scand J Trauma Resusc Emerg Med 2011;19:64.

649. Rizoli S, Nascimento B, Jr., Key N, et al. Disseminated intravascular coagulopathy in the first 24 hours after trauma: the association between ISTH score and anatomopathologic evidence. J Trauma 2011;71:S441-7.

650. Turtay MG, Kirimlioglu V, Ceylan C. Coagulopathy in multiple traumas. Ulus Travma Acil Cerrahi Derg 2010;16:198-202.

651. Windelov NA, Johansson PI, Sorensen AM, et al. Low level of procoagulant platelet microparticles is associated with impaired coagulation and transfusion requirements in trauma patients. J Trauma Acute Care Surg 2014;77:692-700.

652. Chai JK, Liu W, Deng HP, et al. A novel model of burn-blast combined injury and its phasic changes of blood coagulation in rats. Shock 2013;40:297-302.

653. Hayakawa M, Gando S, Ieko M, et al. Massive Amounts of Tissue Factor Induce Fibrinogenolysis without Tissue Hypoperfusion in Rats. Shock 2013;39:514-519.

654. Hayakawa M, Gando S, Ono Y, et al. Noble-Collip Drum Trauma Induces Disseminated Intravascular Coagulation But Not Acute Coagulopathy of Trauma-Shock. Shock 2015;43:261-7.

655. Prat NJ, Montgomery R, Cap AP, et al. Comprehensive evaluation of coagulation in swine subjected to isolated primary blast injury. Shock 2015;43:598-603.

656. Gissel M, Brummel-Ziedins KE, Butenas S, et al. Effects of an acidic environment on coagulation dynamics. J Thromb Haemost 2016;14:2001-2010.

657. Lesperance RN, Lehmann RK, Harold DM, et al. Recombinant factor VIIa is effective at reversing coagulopathy in a lactic acidosis model. J Trauma Acute Care Surg 2012;72:123-9.

658. Martini WZ, Pusateri AE, Uscilowicz JM, et al. Independent contributions of hypothermia and acidosis to coagulopathy in swine. J Trauma 2005;58:1002-9; discussion 1009-10.

659. Martini WZ, Holcomb JB. Acidosis and coagulopathy: the differential effects on fibrinogen synthesis and breakdown in pigs. Ann Surg 2007b;246:831-5.

660. Husari AW, Belzberg H, Kassak K, Dunham CM. Relationship between intensive care complications and costs and initial 24 h events of trauma patients with severe haemorrhage. Emerg Med J 2009;26:340-3.

661. Heinius G, Hahn RG, Sondén A. Hypothermia increases rebleeding during uncontrolled hemorrhage in the rat. Shock 2011;36:60-66.

662. Kheirabadi BS, Crissey JM, Deguzman R, Holcomb JB. In vivo bleeding time and in vitro thrombelastography measurements are better indicators of dilutional hypothermic coagulopathy than prothrombin time. J Trauma 2007;62:1352-9; discussion 1359-61.

663. Kiraly LN, Differding JA, Enomoto TM, et al. Resuscitation with normal saline (NS) vs. lactated ringers (LR) modulates hypercoagulability and leads to increased blood loss in an uncontrolled hemorrhagic shock swine model. The Journal of trauma 2006;61:57-64; discussion 64-5.

664. Pawelczyk NS, Huby Mdel P, Salsbury JR, et al. Association of hemodilution and blood pressure in uncontrolled bleeding. J Surg Res 2013;184:959-65.

665. Wohlauer MV, Moore EE, Droz NM, et al. Hemodilution is not critical in the pathogenesis of the acute coagulopathy of trauma. J Surg Res 2012;173:26-30.

666. de Oliveira Manoel AL, Neto AC, Veigas PV, Rizoli S. Traumatic brain injury associated coagulopathy. Neurocritical care 2015;22:34-44.

667. Greuters S, van den Berg A, Franschman G, et al. Acute and delayed mild coagulopathy are related to outcome in patients with isolated traumatic brain injury. Crit Care 2011;15:R2.

668. Kearney TJ, Bentt L, Grode M, et al. Coagulopathy and catecholamines in severe head injury. J Trauma 1992;32:608-11; discussion 611-2.

Appendix A: Interview Guide and Questions

Clinician Demographics

- Base speciality
- Years of pre-hospital consultancy

Introduction

I would like you to consider your approach to how you asses, and decide treatments for, trauma patients you see in your pre-hospital clinical role. I'm going to ask you about bleeding and trauma induced coagulopathy. Would that be alright? Thinking about bleeding first…

Bleeding

- Is it a fair assumption to suggest you assess a trauma patient for bleeding?
- When do you start thinking about whether or not a patient might be bleeding?
- What do you think are the most important clinical pieces of information you use to assess a patient's severity of bleeding?
- How does hypovolaemic shock change your management for a patient?
- What is your intention when giving blood?
- Can we discuss damage control resuscitation?
- Do you think you practice DCR pre-hospital?
- When do you think DCR becomes necessary pre-hospital?
- What are the components of your Pre-hospital DCR?
- What influences your decision to transfuse pre-hospital blood?
- What is difficult about deciding if a patient needs a transfusion?
- What do you consider a greater risk to a patient: under transfusion of a bleeding patient or unnecessary transfusion to a patient without significant blood loss?

OK thank you. Let's move on to discuss Traumatic Coagulopathy:

Trauma Induced Coagulopathy

- Do you routinely assess for trauma induced coagulopathy?
- What are the most important clinical pieces of information you use to assess a patient's risk of developing coagulopathy?
- What would you say might be difficult for a novice about deciding if a patient is coagulopathic?
- How does your judgement of coagulopathy affect your treatment decisions?
- Thinking about risk thresholds: what do you think is a clinically significant risk for treating coagulopathy? Let me explain what I mean by this question: NELA recommends for a patient with a risk of mortality greater than 10% there ought to be consultant led surgery, ICU post op bed etc… In our series of patients we found that mortality in non-coagulopathic patients is <5%, but is >50% in those who develop coagulopathy. At what risk threshold would you consider changing care?
- What do you consider to be a greater risk to a patient:
  Missed identification of patient with coagulopathy or the potential side effects of haemostatic resuscitation in a patient with normal coagulation.
- How important are these elements of care delayed recognition, no hospital pre-warning, delayed treatment?

Appendix B: EmPHATTIC interview questions

Please answer the following questions on a scale of 1 (no/easy) to 10 (yes/hard) as you believed at the specified timepoint.

1. **Time point 1:** immediately after the initial patient assessment
    - Did you think this patient was coagulopathic?
    - Did you think this patient would need blood in the next 4 hours?

2. **Time point 2:** immediately after handover in the Emergency Department
    - Did you think this patient was coagulopathic?
    - How easy was that prediction?
    - Did you think this patient would need blood in the next 4 hours?
    - Did you think this patient was coagulopathic?
    - How easy was that prediction?
    - Did you think this patient would need blood in the next 4 hours?

3. **Time point 3:** Please answer these questions now you have the model's results
    - Do you agree with the TIC prediction?
    - Do you agree with the transfusion prediction?
    - What (if any) decisions does the model's additional information influence?
    - Do you think this patient was coagulopathic?
    - How easy was that prediction?
    - Do you think this patient would need blood in the next 4 hours?
    - How many PRBC units do you think they will receive in the next 4 hours?
    - Do you think this patient was coagulopathic?
    - How easy is that prediction?

Appendix C: Risk of Bias

Table 31: Risk of bias in human observational studies

Study	Eligibility criteria	Baseline characteristics of sample	Representative sample from a relevant population	All eligible patients included	Coagulopathy appropriately defined	Coagulopathy assessed on arrival to hospital	Causative factors defined	Causative factors measured accurately	Adjustment for confounding	Total Risk of Bias Score
Aucar (2003)	+	-	+	-	+	?	-	-	-	3
Auer (1978)	+	-	+	?	-	-	+	-	-	3
Bredbacka (1994)	+	-	+	?	-	?	+	+	-	4
Brohi (2003)	+	-	+	-	+	+	+	+	-	6
Brohi (2007)	+	+	+	+	-	+	+	+	-	7
Brohi (2008)	+	-	+	+	+	+	+	+	-	7
Brown (2012)	+	-	+	+	+	+	+	+	-	7
Brown (2013)	+	+	+	?	+	+	+	+	+	8
Cap (2011)	+	+	+	+	+	?	+	+	-	7
Carrick (2005)	+	+	+	+	+	?	+	+	-	7
Castelli (2014)	+	+	+	?	-	-	+	+	+	6
Chandler (2010)	-	-	+	?	+	+	-	-	-	3
Charbit (2016)	+	+	+	+	+	+	+	+	-	8
Cheddie (2013)	+	-	+	?	+	+	+	+	-	6
Chhabra (2013)	+	+	+	?	+	?	+	-	-	5
Cohen (2007)	+	+	+	+	+	+	+	+	-	8
Cohen (2009)	+	+	+	+	+	+	-	?	-	6
Cohen (2010)	+	+	+	+	+	+	-	?	-	6
Cohen (2012)	+	+	+	+	+	+	+	+	-	8
Cohen (2013)	+	+	+	+	+	+	+	+	-	8
Cosgriff (1997)	+	+	+	+	+	?	+	+	-	7
Davis (1996)	+	+	+	+	+	?	+	+	-	7
de Oliveira Manoel (2015)	+	+	+	?	+	-	+	-	-	5
Deras (2014)	+	+	+	+	+	+	-	-	-	6
Di Battista (2016)	+	+	+	?	+	+	+	+	+	8
Dunbar (2009)	-	+	+	?	+	+	-	?	-	4
Engels (2011)	+	+	+	?	+	+	+	+	-	7
Ferrara (1990)	+	+	+	+	+	?	+	+	-	7
Floccard (2012)	+	+	+	+	+	+	+	+	-	8
Frith (2010)	+	+	+	-	+	?	+	+	-	6
Genet (2013)	+	+	+	-	+	+	+	+	-	7

Goodnight (1974)	+	+	+	?	-	-	+	-	?	4
Greuters (2011)	+	+	+	?	+	+	+	+	-	7
Halpern (2008)	+	-	+	-	+	+	-	-	-	4
Hayakawa (2011)	+	+	+	?	+	-	+	+	-	6
Hubetamann (2011)	+	+	+	?	-	?	+	+	+	6
Hulka (1996)	+	+	+	?	+	-	+	+	-	6
Husari (2009)	+	+	+	+	-	+	+	+	-	7
Ireland (2011)	+	+	+	?	-	?	+	+	-	5
Jansen (2011)	+	+	+	+	+	+	+	+	-	8
Johansson (2011a)	+	+	+	-	-	+	?	?	-	7
Johansson (2011b)	+	+	+	?	+	+	+	+	-	7
Johansson (2012a)	+	+	+	-	+	+	-	-	-	4
Johansson (2012b)	+	+	+	-	+	+	+	+	-	7
Johansson (2013)	+	+	+	-	+	+	+	+	-	5
Kapsch (1984)	?	?	+	?	?	+	+	?	-	3
Kearney (1992)	+	-	+	+	+	-	+	-	-	5
Keller (2001)	+	-	+	+	+	?	+	+	-	6
Khan (2014)	+	+	+	-	+	+	+	+	-	7
Kutcher (2013)	+	+	+	+	+	+	+	+	-	8
Lee (2014)	+	+	+	+	-	+	+	+	-	7
Lozance (1998)	+	-	+	+	+	-	+	+	-	6
Lustenberger (2010)	+	+	+	+	+	?	+	+	-	7
Lustenberger (2013)	+	+	+	+	+	-	+	+	-	7
MacLeod (2014)	+	+	+	-	+	+	+	+	-	7
Maegele (2007)	+	+	+	+	+	?	+	+	-	7
Matijevic (2014)	+	+	+	?	+	+	+	+	-	7
May (1997)	+	-	+	+	+	+	+	+	-	7
Mitra (2011)	+	+	+	?	+	+	+	+	-	7
Moore (2014)	+	+	+	+	+	-	-	?	-	5
Neal (2014)	+	+	+	-	+	+	+	+	-	7
Nekludov (2007)	+	-	+	?	-	-	+	+	-	4
Niles (2008)	+	+	+	+	+	?	+	+	-	7
Ostrowski (2011)	+	+	+	-	+	+	-	-	-	5
Ostrowski (2012)	+	+	+	-	-	+	+	+	-	6
Peiniger (2012)	+	+	+	?	+	?	+	+	-	6
Raza (2013)	+	+	+	-	+	+	+	+	-	7
Rizoli (2011a)	+	+	+	+	+	+	+	+	-	8
Rizoli (2011b)	+	+	+	-	+	+	-	-	-	5
Schochl (2013)	+	+	+	-	+	+	+	+	-	7
Shaz (2011)	+	+	+	-	+	+	+	+	-	7
Sherren (2013)	+	+	+	+	+	?	+	+	-	7
Simmons (2011)	-	+	+	-	+	?	+	+	+	6
Sixta (2012)	+	+	+	+	+	?	+	+	-	7
Talving (2009)	?	+	+	+	+	?	+	+	-	6
Tonglet (2014)	+	-	+	?	+	+	-	-	-	4
Turtay (2010)	+	-	+	?	-	?	+	-	-	3
van der Sande (1978)	-	-	+	?	-	-	+	+	-	3

Wafaisade (2010a)	+	+	+	+	+	?	+	+	-	7
Wafaisade (2010b)	+	+	+	?	+	?	+	+	-	6
Watts (1998)	+	+	+	+	-	+	+	+	-	7
White (2015a)	+	+	+	?	+	+	+	+	+	8
White (2015b)	+	+	+	?	-	+	-	-	+	5
Windelov (2014)	+	+	+	+	+	+	-	-	-	6
Wu (2014)	+	+	+	-	+	+	+	+	-	7
Xu (2013)	+	+	+	?	+	-	+	+	-	6
Zehtabchi (2008)	+	+	+	-	+	?	+	+	-	6
Lustenberger (2013)	+	+	+	+	+	-	+	+	-	7
MacLeod (2014)	+	+	+	-	+	+	+	+	-	7
Maegele (2007)	+	+	+	+	+	?	+	+	-	7
Matijevic (2014)	+	+	+	?	+	+	+	+	-	7
May (1997)	+	-	+	+	+	+	+	+	-	7
Mitra (2011)	+	+	+	?	+	+	+	+	-	7
Moore (2014)	+	+	+	+	+	-	-	?	-	5
Neal (2014)	+	+	+	-	+	+	+	+	-	7
Nekludov (2007)	+	-	+	?	-	-	+	+	-	4
Niles (2008)	+	+	+	+	+	?	+	+	-	7
Ostrowski (2011)	+	+	+	-	+	+	-	-	-	5
Ostrowski (2012)	+	+	+	-	-	+	+	+	-	6
Peiniger (2012)	+	+	+	?	+	?	+	+	-	6
Raza (2013)	+	+	+	-	+	+	+	+	-	7
Rizoli (2011a)	+	+	+	+	+	+	+	+	-	8
Rizoli (2011b)	+	+	+	-	+	+	-	-	-	5
Schochl (2013)	+	+	+	-	+	+	+	+	-	7
Shaz (2011)	+	+	+	-	+	+	+	+	-	7
Sherren (2013)	+	+	+	+	+	?	+	+	-	7
Simmons (2011)	-	+	+	-	+	?	+	+	+	6
Sixta (2012)	+	+	+	+	+	?	+	+	-	7
Talving (2009)	?	+	+	+	+	?	+	+	-	6
Tonglet (2014)	+	-	+	?	+	+	-	-	-	4
Turtay (2010)	+	-	+	?	-	?	+	-	-	3
van der Sande (1978)	-	-	+	?	-	-	+	+	-	3
Wafaisade (2010a)	+	+	+	+	+	?	+	+	-	7
Watts (1998)	+	+	+	+	-	+	+	+	-	7
White (2015a)	+	+	+	?	+	+	+	+	+	8
White (2015b)	+	+	+	?	-	+	-	-	+	5
Windelov (2014)	+	+	+	+	+	+	-	-	-	6
Wu (2014)	+	+	+	-	+	+	+	+	-	7
Xu (2013)	+	+	+	?	+	-	+	+	+	7
Zehtabchi (2008)	+	+	+	-	+	?	+	+	-	6

+ low risk of bias; - high risk of bias; ? unclear risk of bias;

Table 32: Risk of bias in human experimental studies

Study	Appropriate / Random sequence generation	Allocation concealment / homogeneous sample	Blinding / automated	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other bias	Total Risk of Bias Score
Bolliger, 2010	+	+	+	+	+	+	+	7
Caballo, 2013	+	+	+	+	+	+	+	7
Coats, 2006	+	+	+	+	+	+	+	7
Darlington, 2012	+	+	+	+	+	+	-	6
Delano, 2015	+	+	+	+	+	+	?	6
Dirkmann, 2013	+	+	+	+	+	+	+	7
Engstrom, 2006	+	+	+	+	+	+	+	7
Evans, 1996	?	?	?	+	?	+	+	3
Gissel, 2016	+	+	+	+	+	-	+	6
Howard, 2016	+	+	+	+	+	+	+	7
Kettner, 2003	-	-	-	-	?	+	+	2
Schols, 2008	+	+	+	-	+	+	?	5
Smith, 2015	+	+	+	+	+	?	-	5
Sossdorf, 2009	+	+	+	+	+	+	+	7
Whelihan, 2014	+	+	+	+	+	+	+	7
Wolberg, 2004	+	+	+	+	+	NR	+	6

+ low risk of bias; - high risk of bias; ? unclear risk of bias; NR, Not reported.

Table 33: Risk of bias in animal experimental studies

Study	Random sequence generation	Baseline characteristics	Allocation concealment	Blinding	Blinding of outcome assessment	Incomplete outcome data	Other bias	Total Risk of Bias Score
Castellino (2014)	n/a	-	n/a	n/a	?	?	?	0
Chai (2013)	+	+	+	-	+	+	+	6
Chen (2016)	?	-	?	-	+	+	?	2
Chesebro (2009)	?	+	?	?	+	+	+	4
Cho (2009)	n/a	+	n/a	n/a	+	+	+	4
Darlington (2011)	-	+	n/a	-	?	+	+	3
Darlington (2013)	n/a	+	n/a	n/a	+	?	+	3
Darlington (2015)	+	+	?	-	+	+	+	5
Dickneite (2009)	+	+	?	?	+	+	-	4
Dickneite (2010)	-	+	-	?	+	+	-	3
Donahue (2014)	n/a	+	n/a	n/a	+	?	+	3
Doran (2012)	+	+	?	-	+	+	+	5
Duan (2014)	-	+	-	-	+	?	+	3
Fries (2005)	+	+	?	+	+	+	-	5
Fries (2006)	+	+	?	+	+	+	-	5
Frith (2010)	+	+	?	-	+	+	+	5
Fung (2013)	-	+	-	-	+	+	+	4
Grottke (2010)	+	+	+	+	+	+	+	7
Hagemo (2013)	+	+	?	-	+	+	+	5
Harr (2011)	n/a	+	n/a	n/a	+	+	+	4
Hayakawa (2013)	+	+	?	-	+	+	+	5
Hayakawa (2015)	+	+	?	-	+	+	+	5
Heinius (2011)	+	+	?	-	-	+	+	4
Iwamoto (2010)	+	+	?	-	+	+	+	5
Kheirabadi (2007)	+	+	?	-	?	+	+	4
Kiraly (2006)	+	+	+	+	+	+	+	7
Klemcke (2005)	+	+	?	+	+	+	+	6
Lesperance (2012)	+	+	?	-	+	?	+	4
Letson (2012)	+	-	?	-	+	+	-	3
Letson (2016)	+	+	?	-	+	+	+	5
Martini (2005a)	+	+	?	-	+	+	+	5
Martini (2005b)	+	+	?	-	+	+	+	5
Martini (2006)	+	+	?	-	+	+	+	5
Martini (2006)	+	+	?	-	+	+	+	5
Martini (2007)	+	+	?	-	+	+	+	5

Martini (2008)	+	+	?	-	+	+	+	5
Martini (2013)	+	+	?	-	+	+	+	5
Mohr (2013)	+	+	?	-	+	+	+	5
Nishi (2013)	+	+	?	-	?	+	+	4
Park (2013)	+	+	+	-	+	+	+	6
Pawelczyk (2013)	+	+	?	-	+	+	+	5
Pragst (2010)	+	+	?	+	+	?	-	4
Prat (2015)	n/a	+	n/a	n/a	n/a	-	+	2
Rezende-Neto (2014)	+	+	?	-	+	+	+	5
Spronk (2015)	+	+	+	+	+	+	+	7
Tian (2015)	n/a	-	n/a	n/a	+	?	+	2
Torres (2013)	+	+	?	-	+	+	+	5
van Zyl (2016)	?	+	?	-	+	+	+	4
Watts (2015)	+	+	?	-	+	+	+	5
White (2010)	n/a	+	n/a	n/a	+	+	+	4
White (2011)	n/a	+	n/a	n/a	+	+	+	4
Wohlauer (2012)	n/a	+	n/a	n/a	n/a	-	-	1
Xu (2015)	+	+	?	-	+	+	+	5
Xu (2015)	+	+	?	-	+	+	+	5

+ low risk of bias; - high risk of bias; ? unclear risk of bias; n/a, Not Applicable.

Appendix D: Synthesis of Supporting Evidence

Table 34: Evidence supporting a causal relationship between Hypoperfusion and coagulopathy

Author, Year	Design	Study Population / Sample	Inclusion Criteria / Exposure	Sample Size	Coagulation measurement	Association	No confounding	**Direct** Temporality	Biological Grad.	Reversibility	**Mechanistic** P. mechanism	**Parallel** Replicability	Similarity
Human Observational Studies													
Brohi, 2007[449]	PC	SC, Civilian, all MOI	Major trauma	208	PT, PPT	+	-	+	+	0	+		
Brohi, 2008[451]	PC	SC, Civilian, all MOI	Major trauma	208	PT, PPT	+	-	+	+	0	+		
Brown, 2012[619]	PC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1877	INR	+	-	+	0	0	0		
Brown, 2013[620]	PC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI ISS>15	1216	INR	+	-	+	0	0	0		
Cap, 2011[452]	RC	MC, Military, all MOI	iTBI	1609	INR	+	-	+	+	0	0		

Charbit, 2016[621]	RC	SC, Civilian, all MOI	Admitted to trauma ICU	704	PTr, aPPTr	+	-	+	0	0	0
Cheddie, 2013[455]	RC	SC, Civilian, all MOI	BD >-2	28	INR	+	-	+	0	0	0
Cohen, 2007[450]	PC	SC, Civilian, all MOI	iTBI	39	PT, PPT	+	+ [a]	+	0	0	0
Cohen, 2009[622]	PC	SC, Civilian, all MOI	All trauma patients	168	INR	+	+ [a]	+	+	0	0
Cohen, 2012[54]	PC	SC, Civilian, all MOI	Highest TA	203	INR	+	-	+	+	0	0
Cohen, 2013[440]	PC	MC, Civilian, all MOI	Blood transfusion within 6hrs	1198	INR, PTT	+	-	+	+	0	0
Cosgriff, 1997[255]	PC	SC, Civilian, all MOI	>10u PRBC/24hr	58	PT, PTT	+	-	+	+	0	0
Davis, 1996[448]	RC	SC, Civilian, all MOI	ABG within 1hr of arrival	2954	PT, PTT, Plats, Clinical	+	-	+	+	0	0
Deras, 2014[457]	RC	SC, Civilian, all MOI	Admitted to ICU	663	PT, aPPT	+	-	+	0	0	0
Ferrara, 1990[500]	RC	SC, Civilian, all MOI	>12u PRBC/24hr, without head injury	45	Clinical	+	-	+	+	0	0
Frith, 2010[395]	RC	MC, Civilian, all MOI	TA	3646	PTr	+	-	+	+	0	0
Hayakawa, 2011[623]	PC	SC, Civilian, all MOI	ISS>8, survived >24hrs	57	PT, Plats	+	-	+	0	0	0
Jansen, 2011[470]	RC	SC, Civilian, all MOI	ISS>15, within 2hrs of injury	71	INR	+	-	+	+	0	+
Johansson, 2012b[624]	PC	SC, Civilian, all MOI	Full TA	80	INR, aPPT, ROTEM	+	-	+	+	0	+
Kapsch, 1984[447]	PC	SC, Civilian, all MOI	Within 4hrs of injury	23	PT, aPTT, euglobulin fibrinolysis, Plats	+	-	+	+	0	+
Kutcher, 2013[625]	PC	SC, Civilian, all MOI	Highest TA	163	INR, PPT	+	-	+	0	0	0
Lee, 2014[458]	PC	MC, Civilian, all MOI	Highest TA, AIS>2 in at least one region,	462	INR, TEG	+	-	+	+	0	0

Lustenberger, 2010[495]	RC	MC, Civilian, all MOI	iTBI admitted to ICU	132	INR, aPTT, Plats	+	-	+	+	0	0
Matijevic, 2014[478]	CC	MC, Civilian, all MOI	Blood transfusion within 6hrs	180	INR, aPTT	+	-	+	+	0	+
Mitra, 2011[257]	RC & PC	SC, Civilian, all MOI	ISS > 15, or urgent surgery, or ICU or death	2905	INR, aPTT	+	+ [a]	+	+	0	0
Neal, 2014[503]	RC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1897	INR, clinical	+	+ [a]	+	0	0	0
Raza, 2013[445]	PC	SC, Civilian, all MOI	TA	288	ROTEM	+	-	+	+	0	+
Schochl, 2013[487]	PC	SC, Civilian, non-trauma	Out of hospital cardiac arrest receiving CPR	53	ROTEM	+	-	+	+	0	0
Simmons, 2011[453]	RC	MC, Military, GSW or Explosive MOI	>9 PRBCs within 24hrs,	450	INR	+	-	+	0	0	0
Sixta, 2012[454]	RC	SC, Civilian, all MOI	Immediate laparotomy	1218	INR	+	+ [a]	+	+	0	0
Talving, 2009[441]	PC	SC, Civilian, all MOI	Admitted to ICU, AIS head >2	436	INR, aPPT, Plats	+	+ [a]	+	+	0	0
Tonglet, 2014[262]	PC	SC, Civilian, all MOI	Severe trauma	82	INR, ROTEM	+	-	+	+	0	0
Wafaisade, 2010a[566]	RC	MC, Civilian, Blunt trauma	German Trauma Registry, iTBI	3114	PT, Plats	+	+ [a]	+	0	0	0
Wafaisade, 2010b[442]	RC	MC, Civilian, all MOI	German Trauma Registry, ISS >15	1987	PT, Plats	+	+ [a]	+	0	0	0
White, 2015a[626]	CC	SC, Civilian, all MOI	admitted, survived >36hrs	95	INR, aPPT, TEG	+	-	+	+	0	+
White, 2015b[459]	PC	SC, Civilian, all MOI	TA, admitted, survived >36hrs	84	INR, aPPT, TEG	+	-	+	+	0	+
Xu, 2013[456]	RC	SC, Civilian, all MOI	ISS>15, admitted to ICU <24hrs of injury	223	INR, PPT, aPPT	+	+ [a]	+	+	0	0
Human Experimental Studies											
Nil											

Animal Experimental Studies

Chen, 2017[627]	Controlled Trial	Rat	Controlled haemorrhage to MAP 40mmHg	40	PT, aPTT, ROTEM	+	-	+	0	0	+	+	+
Chesebro, 2009[66]	Controlled Trial	Mouse	Controlled haemorrhage to MAP 35mmHg	40	aPTT	+	+	+	0	0	+	+	+
Cho, 2009[628]	AMD	Swine	Controlled 60% TBV haemorrhage	37	PT, TEG	+	-	+	0	0	0	+	+
Darlington, 2011[504]	AMD	Swine	Controlled haemorrhage to MAP 30mmHg	18	PT, aPPT, TEG	+	+	+	0	-	0	0	+
Darlington, 2013[468]	AMD	Rat	Controlled 40% TBV haemorrhage	50	PT, aPTT, ROTEM, Plats	+	- [b]	+	+	0	0	+	+
Darlington, 2015[629]	RCT	Rat	Controlled 40% TBV haemorrhage	63	PT, aPTT, ROTEM	+	- [b]	+	0	0	+	+	+
Doran, 2012[630]	RCT	Swine	Controlled 35%TBV haemorrhage + grade 4 liver injury	24	PT	+	-	+	0	0	0	+	+
Duan, 2014[631]	Controlled Trial	Swine	Controlled haemorrhage to MAP 40mmHg	22	INR, PT, ROTEM	+	- [b]	+	0	0	+	+	+
Frith, 2010[395]	RCT	Rat	Controlled haemorrhage to MAP 40050mmHg	40	PT, aPTT	+	+	+	0	0	0	+	+
Fung, 2013[632]	Controlled Trial	Ovine	Controlled 35% TBV haemorrhage	14	PT, aPPT, ROTEM	+	+	+	0	0	0	0	0
Hagemo, 2013[633]	RCT	Swine	Controlled 45% TBV haemorrhage	18	INR, ROTEM	+	+	+	0	0	0	0	+
Harr, 2011[462]	No control group	Rat	Controlled haemorrhage to MAP 35mmHg	6	ROTEM	+	- [b]	+	0	0	+	+	+
Iwamoto, 2010[634]	RCT	Rat	Controlled haemorrhage	24	Sonoclot	+	+	+	0	0	0	+	+
Letson, 2012[466]	RCT	Rat	Controlled haemorrhage to 35040mmHg	68	PT, aPPT	+	+	+	+	+	0	+	+
Letson, 2016[477]	RCT	Rat	Controlled haemorrhage to 35040mmHg	30	PT, aPPT, ROTEM	+	+	+	+	0	+	+	+
Martini, 2005[460]	RCT	Swine	Controlled 35% TBV haemorrhage	12	TEG	0	+	+	0	0	+	+	+

Martini, 2006b[635]	RCT	Swine	Controlled 35% TBV haemorrhage	18		+	+	+	0	0	+	+	+
Martini, 2008[520]	RCT	Swine	Controlled 35% TBV haemorrhage	24	PT, aPTT, TEG, ACT	+	-	+	0	0	+	+	+
Martini, 2013[544]	RCT	Swine	Controlled 60% TBV haemorrhage	21	TEG	+	-	+	0	+	+	+	+
Nishi, 2013[636]	RCT	Rat	Controlled haemorrhage + tail amputation	18	Sonoclot	+	-	+	+	0	0	+	+
Park, 2013[467]	RCT	Rat	Uncontrolled haemorrhage 0 Grade 3 splenic injury	32	ROTEM	+	+	+	0	0	0	+	+
Rezende-Neto, 2014[481]	RCT	Rat	Controlled 40% TBV haemorrhage	24	ROTEM	+	+	+	0	0	+	+	+
Torres, 2013[545]	RCT	Rat	Controlled 40% TBV haemorrhage	32	ROTEM	+	-	+	+	0	+	+	+
van Zyl, 2016[446]	AMD	Ovine	Controlled haemorrhage of 20% TBV (moderate) and 30% TBV (severe)	12	PT, aPTT, ROTEM	+	- [b]	+	+	0	+	+	+
Watts, 2015[539]	RCT	Swine	Controlled 35% TBV haemorrhage	24	PT, aPTT, TEG	+	-	+	0	0	+	+	+
White, 2010[461]	Controlled Trial	Swine	Controlled haemorrhage to MAP 30mmHg	23	PT, aPTT, TEG	+	- [b]	+	0	0	0	+	+
White, 2011[485]	No control group	Swine	Controlled haemorrhage to MAP 30mmHg	17	PT, PTT, TEG	+	- [b]	+	0	0	+	0	+
Xu, 2015a[637]	RCT	Rat	Controlled haemorrhage to MAP 35mmHg	25	PT, aPTT	+	- [b]	+	0	0	+	+	+
Xu, 2015b[638]	RCT	Rat	Controlled haemorrhage to MAP 30mmHg	152	PT, aPTT	+	- [b]	+	+	0	+	+	+

[a] Multivariable adjustment for confounding, [b] Combined tissue injury and haemorrhage model.

ACT, Activated Clotting Time; AIS, Abbreviated Injury Scale; AMD, Animal model development; BD, Base Deficit; CC, Case-Control; IV, In Vitro; CPR, Cardio pulmonary resuscitation; GSW, Gunshot wound; HS, Haemorrhagic Shock; INR, International Normalised Ratio; ISS, Injury severity score; iTBI, isolated Traumatic Brain Injury; MOI, Mechanism of Injury; NS, Normal Saline; PC, Prospective Cohort; PH, Pre-hospital; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; TA, Trauma Activation; TBV, Total Blood Volume; TI, Tissue Injury; TT, Thrombin Time.

Table 35: Evidence supporting a causal relationship between tissue injury and coagulopathy

								Direct			**Mechanistic**	**Parallel**	
Author, Year	**Design**	**Study Population / Sample**	**Inclusion Criteria / Exposure**	**Sample Size**	**Coagulation measurement**	Association	No confounding	Temporality	Biological Grad.	Reversibility	P. mechanism	Replicability	Similarity
Human Observational Studies													
Brohi, 2003[55]	RC	SC, Civilian, all MOI	Admitted by Helicopter emergency service	1867	PT, aPTT. TT	+	-	+	+	0	0		
Brown, 2012[619]	PC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1877	INR	+	-	+	0	0	0		
Cap, 2011[452]	RC	MC, Military, all MOI	iTBI,	1609	INR	+	+	+	+	0	0		
Carrick, 2005[557]	RC	SC, Civilian, Blunt Trauma	iTBI, GCS <14	184	PT, PPT, Plats	+	-	+	+	0	0		
Cheddie, 2013[455]	RC	SC, Civilian, all MOI	BD >-2	28	INR	+	-	+	0	0	0		
Chhabra, 2013[568]	PR	SC, Civilian, all MOI	iTBI, GCS <13, admitted neurosurgery	208	PT, aPPT	+	-	+	+	0	0		
Cohen, 2009[622]	PC	SC, Civilian, all MOI	TA	168	INR	+	-	+	+	0	0		
Cohen, 2010[639]	PC	SC, Civilian, all MOI	TA	168	INR	+	-	+	0	0	+		

Cohen, 2012[54]	PC	SC, Civilian, all MOI	Highest TA	203	INR	+	-	+	+	0	0
Cohen, 2013[440]	PC	MC, Civilian, all MOI	Blood transfusion within 6hrs	1198	INR, PTT	+	+ [a]	+	+	0	0
Cosgriff, 1997[255]	PC	SC, Civilian, all MOI	>10u PRBC/24hr	58	PT, PTT	+	+ [a]	+	+	0	0
Davis, 1996[448]	RC	SC, Civilian, all MOI	ABG within 1hr of arrival	2954	PT, PTT, Plats, Clinical	+	-	+	+	0	0
Deras, 2014[457]	RC	SC, Civilian, all MOI	Admitted to ICU	663	PT, aPPT	+	-	+	0	0	0
Dunbar, 2009[640]	CC	SC, Civilian, all MOI	Blood sample within 1hr of ED arrival	42	INR, PT	+	-	0	0	0	+
Floccard, 2012[493]	PC	SC, Civilian, Blunt MOI	Attend by SAMU, pre-hospital blood sample	45	PT, Plats	+	-	+	+	0	0
Frith, 2010[395]	RC	MC, Civilian, all MOI	TA	3646	PTr	+	-	+	+	0	0
Genet, 2013[494]	RC	SC, Civilian, all MOI	Full TA	80	INR, aPPT	+	-	+	+	0	0
Hulka, 1996[556]	RC	SC, Civilian, Blunt MOI	Had CT brain, blood sample within 24hrs	159	PT, PPT, Plats	+	-	+	0	0	0
Johansson, 2011a[471]	RC	SC, Civilian, all MOI	Full TA	80	INR, aPPT	+	-	+	+	0	0
Johansson, 2011b[641]	RC	SC, Civilian, all MOI	Full TA	75	INR, aPPT	+	-	+	0	0	0
Johansson, 2012[642]	PC	SC, Civilian, all MOI	Full TA	80	INR, aPPT, ROTEM	+	-	+	+	0	+
Johansson, 2013[643]	PC	SC, Civilian, all MOI	Full TA	80	INR, aPPT	+	-	+	+	0	+
Kapsch, 1984[447]	PC	SC, Civilian, all MOI	Within 4hrs of injury	23	PT, aPTT, euglobulin fibrinolysis, Plats	+	-	+	+	0	+
Khan, 2014[644]	RC	MC, Civilian, all MOI	TA, received >3unit PRBC	106	ROTEM, PTr	+	-	+	0	0	0
Kutcher, 2013[625]	PC	SC, Civilian, all MOI	Highest TA	163	INR, PPT	+	-	+	0	0	0

Lee, 2014[458]	PC	MC, Civilian, all MOI	Highest TA, AIS>2 in at least one region,	462	INR, TEG	+	-	+	+	0	0
Lozance, 1998[563]	RC	SC, Civilian, all MOI	iTBI	105	PT, PPT, TT, Plats	+	-	+	+	0	0
Lustenberger, 2010[495]	RC	MC, Civilian, all MOI	iTBI admitted to ICU	132	INR, aPTT, Plats	+	-	+	+	0	0
Lustenberger, 2013[645]	PC	SC, Civilian, all MOI	ISS >15	26	INR, aPTT, Plats	+	-	+	0	0	+
MacLeod, 2014[646]	PC	SC, Civilian, all MOI	TA, admitted for >24hrs	701	PT	+	-	+	+	0	0
Maegele, 2007[528]	RC	MC, Civilian, all MOI	Entry into German Trauma Registry	8724	Quick Test, Plats	+	-	+	+	0	0
Matijevic, 2014[478]	CC	MC, Civilian, all MOI	Blood transfusion within 6hrs	180	INR, aPTT	+	-	+	+	0	+
May, 1997[647]	RC	SC, Civilian, Blunt Trauma	iTBI, GCS <9	26	Pt, PPT	+	-	+	+	0	0
Mitra, 2011[257]	RC & PC	SC, Civilian, all MOI	ISS > 15, or urgent surgery, or ICU or death	2905	INR, aPTT	+	+ [a]	+	+	0	0
Moore, 2014[188]	PC	SC, Civilian, all MOI	ISS >15	180	INR, TEG	+	-	+	0	0	0
Neal, 2014[503]	RC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1897	INR, clinical	+	+ [a]	+	0	0	0
Nekludov, 2007[573]	PC	SC, Civilian, Blunt Trauma	iTBI, GCS<9, admitted to neuroICU	11	INR, bleeding time, Plats	+	-	+	0	0	0
Niles, 2008[490]	RC	SC, Military, all MOI	Received a blood transfusion	391	INR	+	-	+	+	0	0
Ostrowski, 2011[648]	PC	SC, Civilian, all MOI	Full TA	80	INR, aPPT, TEG	+	-	+	+	0	+
Ostrowski, 2012[81]	PC	SC, Civilian, all MOI	Full TA	77	INR, aPPT, TEG	+	-	+	0	0	0
Raza, 2013[445]	PC	SC, Civilian, all MOI	TA	288	ROTEM	+	-	+	+	0	+
Rizoli, 2011a[69]	PC	SC, Civilian, all MOI	ISS>15, no PH transfusion	110	INR, aPTT	+	-	+	0	0	0

Rizoli, 2011b[649]	PC	SC, Civilian, all MOI	ISS>15, no PH transfusion	423	PT, Plats	+	-	+	0	0	0		
Shaz, 2011[492]	CC	SC, Civilian, all MOI	TA	91	PT	+	-	+	0	0	0		
Simmons, 2011[453]	RC	MC, Military, GSW or Explosive MOI	>9 PRBCs within 24hrs,	450	INR	+	-	+	0	0	0		
Sixta, 2012[454]	RC	SC, Civilian, all MOI	Immediate laparotomy	1218	INR	+	-	+	+	0	0		
Talving, 2009[441]	PC	SC, Civilian, all MOI	Admitted to ICU, AIS head >2	436	INR, aPPT, Plats	+	+ [a]	+	+	0	0		
Tonglet, 2014[262]	PC	SC, Civilian, all MOI	Severe trauma	82	INR, ROTEM	+	-	+	+	0	0		
Turtay, 2010[650]	PC	SC, Civilian, all MOI	Unclear	50	INR, aPPT	+	-	+	+	0	0		
Wafaisade, 2010b[442]	RC	MC, Civilian, all MOI	German trauma registry, ISS>15	1987	PT, Plats	+	+ [a]	+	0	0	0		
White, 2015a[626]	CC	SC, Civilian, all MOI	admitted, survived >36hrs	95	INR, aPPT, TEG	+	-	+	+	0	+		
White, 2015b[459]	PC	SC, Civilian, all MOI	TA, admitted, survived >36hrs	84	INR, aPPT, TEG	+	-	+	+	0	+		
Windelov, 2014[651]	PC	SC, Cilivian, all MOI	TA, arterial cannula	210	ROTEM	+	-	+	0	0	0		
Xu, 2013[456]	RC	SC, Civilian, all MOI	ISS>15, admitted to ICU <24hrs of injury	223	INR, PPT, aPPT	+	-	+	+	0	0		
Human Experimental Studies													
Nil													
Animal Experimental Studies													
Chai, 2013[652]	RCT	Rat	Blast and full thickness burn: 25% TBSA	256	PT, aPPT	+	-	+	0	0	0	+	+
Chen, 2017[627]	Controlled Trial	Rat	Injury to bowel, liver, muscle and femur fracture	40	PT, aPTT, ROTEM	+	- [b]	+	0	0	+	+	+

Chesebro, 2009[66]	Controlled Trial	Mouse	Laparotomy	40	aPTT	+	- [b]	+	0	0	+	+	+
Cho, 2009[628]	AMD	Swine	Femur fracture	37	PT, TEG	+	- [b]	+	0	0	0	+	+
Darlington, 2013[468]	AMD	Swine	Injury to bowel, liver, muscle, femur and laparotomy	18	PT, aPPT, TEG	+	- [b]	+	+	0	0	+	+
Darlington, 2015[629]	AMD	Rat	Injury to bowel, liver, muscle and femur fracture	50	PT, aPTT, ROTEM, Plats	+	- [b]	+	0	0	+	+	+
Duan, 2014[631]	Controlled Trial	Swine	Injury to bowel, liver, femur and laparotomy	22	INR, PT, ROTEM	+	- [b]	+	0	0	+	+	+
Frith, 2010[395]	RCT	Rat	Laparotomy and bilateral lower limb fractures	40	PT, aPTT	+	+	+	0	0	0	+	+
Grottke, 2010[538]	RCT	Swine	Grade 3 blunt liver injury	18	PT, aPTT, ROTEM	+	- [b]	+	0	0	+	+	+
Hagemo, 2013[633]	RCT	Swine	Bilateral thigh high-energy GSW (femoral fractures and soft tissue injury)	18	INR, ROTEM	+	- [b]	+	0	0	0	0	+
Harr, 2011[462]	No control group	Rat	Laparotomy	6	ROTEM	+	- [b]	+	0	0	+	+	+
Hayakawa, 2013[653]	RCT	Rat	Exogenous Tissue Factor	18	PT	+	+	+	+	0	+	0	+
Hayakawa, 2015[654]	RCT	Rat	Noble-Collip drum	18	PT	+	-	+	0	0	+	+	+
Klemcke, 2005[549]	RCT	Swine	Grade 5 liver injury	54	PT, aPPT, TEG	+	- [b]	+	+	+	+	0	+
Prat, 2015[655]	AMD	Swine	Isolated blast: 400-kPa peak overpressure	13	PT, aPTT, ROTEM	0	-	0	0	0	0	0	+
Spronk, 2015[540]	RCT	Swine	Grade 3 blunt liver injury	28	PT, aPTT, ROTEM	+	- [b]	+	0	+	+	+	+
van Zyl, 2016[446]	AMD	Ovine	Bilateral tibia fractures and pulmonary contusions (moderate), additional crush injury (severe)	12	PT, aPTT, ROTEM	+	- [b]	+	+	0	+	+	+
Watts, 2015[539]	RCT	Swine	Muscle contusion, laparotomy	24	PT, aPTT, TEG	+	- [b]	+	0	0	+	+	+

White, 2010[461]	Controlled Trial	Swine	Femur fracture, soft tissue injury	23	PT, aPTT, TEG	+	- [b]	+	0	0	0	+	+
White, 2011[485]	No control group	Swine	Femur fracture, soft tissue injury	17	PT, PTT, TEG	+	- [b]	+	0	0	+	0	+
Xu, 2015a[637]	RCT	Rat	Laparotomy	25	PT, aPTT	+	- [b]	+	0	0	+	+	+
Xu, 2015b[638]	RCT	Rat	Laparotomy and bilateral femur fractures	152	PT, aPTT	+	- [b]	+	+	0	+	+	+

[a] Multivariable adjustment for confounding. [b] Combined tissue injury and haemorrhage model.

ACT, Activated Clotting Time; AIS, Abbreviated Injury Scale; AMD, Animal model development; BD, Base Deficit; CC, Case-Control; IV, In Vitro; CPR, Cardio pulmonary resuscitation; GSW, Gunshot wound; HS, Haemorrhagic Shock; INR, International Normalised Ratio; ISS, Injury severity score; iTBI, isolated Traumatic Brain Injury; MOI, Mechanism of Injury; NS, Normal Saline; PC, Prospective Cohort; PH, Pre-hospital; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; TA, Trauma Activation; TBV, Total Blood Volume; TI, Tissue Injury; TT, Thrombin Time.

Table 36: Evidence supporting a causal relationship between acidaemia and coagulopathy

Author, Year	Design	Study Population / Sample	Inclusion Criteria / Exposure	Sample Size	Coagulation measurement	Association	No confounding	Direct: Temporality	Direct: Biological Grad.	Direct: Reversibility	Mechanistic: P. mechanism	Parallel: Replicability	Parallel: Similarity
Human Observational Studies													
Aucar, 2003[501]	PC	SC, Civilian, all MOI	Urgent surgical intervention; chest, abdo or extremity	31	Clinical	+	-	+	0	0	0		
Cosgriff, 1997[255]	PC	SC, Civilian, all MOI	>10u PRBC/24hr	58	PT, PTT	+	+ [a]	+	0	0	0		
Davis, 1996[448]	RC	SC, Civilian, all MOI	ABG within 1hr of arrival	2954	PT, PTT, Plats, clinical	+	-	+	+	0	0		
Engels, 2011[502]	RC	SC, Civilian, all MOI	INR>1.3, arrived within 1hr of injury	290	INR	+	+ [a]	+	+	0	0		
Ferrara, 1990[500]	RC	SC, Civilian, all MOI	>12u PRBC/24hr, without TBI	45	Clinical	+	-	+	0	0	0		
Kashuk, 2010[515]	RC	SC, Civilian, all MOI	Required transfusion	61	INR, PTT, TEG	+	-	+	+	0	0		
Neal, 2014[503]	RC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1897	INR, clinical	+	-	+	0	0	0		
Niles, 2008[490]	RC	SC, Military, all MOI	Received a blood transfusion	391	INR	+	-	+	+	0	0		
Rizoli, 2011a[69]	PC	SC, Civilian, all MOI	ISS>15, no PH transfusion,	110	INR, aPTT	+	-	+	0	0	0		

Human Experimental Studies													
Dirkmann, 2013[508]	IV	Healthy whole blood	pH 7.1 and 6.9	10	ROTEM	+	+	+	+	0	+	+	0
Engstrom, 2006[507]	IV	Healthy whole blood	pH 7.2, 7.0, 6.8	6	ROTEM	+	+	+	+	+	+	+	+
Gissel, 2016[656]	IV	Healthy whole blood	pH 7.0	7	ROTEM	+	+	+	0	0	+	+	0
Animal Experimental Studies													
Cho, 2009[628]	AMD	Swine	Multi-modal coagulopathy model	37	PT, TEG	+	-	+	0	0	0	+	+
Darlington, 2011[504]	AMD	Swine	pH 7.1	18	PT, aPPT, TEG	+	+	+	0	0	0	0	+
Lesperance, 2012[657]	RCT	Swine	Lactic acidosis model	10	INR, ROTEM	+	-	+	0	-	0	+	+
Martini, 2005b[658]	RCT	Swine	pH 7.1	24	PT, Bleeding time	+	+	+	0	0	+	+	+
Martini, 2006a[506]	RCT	Swine	pH 7.1	18	PT, PTT, TEG	+	+	+	0	-	+	+	+
Martini, 2007b[659]	RCT	Swine	pH 7.1	12	PT, PTT, ACT, TEG	+	+	+	0	0	+	+	+

[a] Multivariable adjustment for confounding.

ACT, Activated Clotting Time; AIS, Abbreviated Injury Scale; AMD, Animal model development; BD, Base Deficit; CC, Case-Control; IV, In Vitro; CPR, Cardio pulmonary resuscitation; GSW, Gunshot wound; HS, Haemorrhagic Shock; INR, International Normalised Ratio; ISS, Injury severity score; iTBI, isolated Traumatic Brain Injury; MOI, Mechanism of Injury; NS, Normal Saline; PC, Prospective Cohort; PH, Pre-hospital; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; TA, Trauma Activation; TBV, Total Blood Volume; TI, Tissue Injury; TT, Thrombin Time.

Table 37: Evidence supporting a causal relationship between hypothermia and coagulopathy

Author, Year	Design	Study Population / Sample	Inclusion Criteria / Exposure	Sample Size	Coagulation measurement	Association	No confounding	**Direct** Temporality	Biological Grad.	Reversibility	**Mechanistic** P. mechanism	Replicability	**Parallel** Similarity
Human Observational Studies													
Cohen, 2013[440]	PC	MC, Civilian, all MOI	Blood transfusion within 6hrs	1198	INR, PTT	+	+ [a]	+	+	0	0		
Cosgriff, 1997[255]	PC	SC, Civilian, all MOI	>10u PRBC/24hr	58	PT, PTT	+	+ [a]	+	+	0	0		
Engels, 2011[502]	RC	SC, Civilian, all MOI	Arrived within 1 hr of injury	2473	INR	+	+ [a]	+	+	0	0		
Ferrara, 1990[500]	RC	SC, Civilian, all MOI	>12u PRBC/24hr, without head injury	45	Clinical	+	-	+	0	0	0		
Husari, 2009[660]	PC	SC, Civilian, all MOI	SBP <90, >2L of fluids <1hr, 4U PRBCs <6hrs	27	PT, PPT, Plats	+	-	+	0	0	0		
Ireland, 2011[513]	RC	SC, Civilian, all MOI	Major Trauma, temp within 15mins ED arrival	732	INR, PT, aPPT	+	-	+	0	0	0		
Mitra, 2011[257]	RC & PC	SC, Civilian, all MOI	ISS > 15, or urgent surgery, or ICU or death	2905	INR, aPTT	+	+ [a]	+	+	0	0		
Wafaisade, 2010b[442]	RC	MC, Civilian, all MOI	German trauma registry, ISS>15	1987	PT, Plats	+	-	+	0	0	0		
Watts, 1998[526]	PC	SC, Civilian, all MOI	TA, ISS > 8	112	TEG	+	+ [a]	0	+	0	+		

Xu, 2013[456]	RC	SC, Civilian, all MOI	ISS>15, admitted to ICU <24hrs of injury	223	INR, PPT, aPPT	+	-	+	0	0	0		
Human Experimental Studies													
Darlington, 2012[464]	IV	Healthy whole blood	34ºC	9	TEG, PT, aPTT	+	+	+	0	0	0	+	+
Dirkmann, 2013[508]	IV	Healthy whole blood	33ºC	10	ROTEM	+	+	+	0	0	+	+	+
Kettner, 2003[516]	No control group, in-Vivo	Elective Intracranial surgery patients	36ºC, 34ºC, 32ºC	16	PT, aPPT, TEG, Plats, closure time	+	+	+	+	0	0	0	+
Whelihan, 2014[517]	IV	Healthy whole blood	27ºC	8	TEG	+	+	+	+	0	+	+	+
Wolberg, 2004[518]	IV	Healthy plasma	35ºC, 33ºC, 29ºC, 25ºC and 23ºC	NR	aPTT	+	+	+	+	0	+	+	+
Animal Experimental Studies													
Cho, 2009[628]	AMD	Swine	33ºC	37	PT, TEG	+	-	+	0	0	0	+	+
Heinius, 2011[661]	RCT	Rat	30ºC for 1 hour	40	Clinical	+	+	+	0	0	0	+	+
Iwamoto, 2010[634]	RCT	Rat	33ºC	24	Sonoclot	+	+	+	0	0	0	+	+
Kheirabadi, 2007[662]	RCT	Rabbit	34ºC	14	PT, aPPT, BT, TEG	+	-	+	0	0	0	+	+
Klemcke, 2005[549]	RCT	Swine	32.5ºC	54	PT, aPTT, TEG	+	-	+	0	0	0	0	+
Martini, 2005b[658]	RCT	Swine	32ºC	24	PT, BT	+	+	+	0	0	+	+	+
Martini, 2007[510]	RCT	Swine	32ºC	12	TEG	+	+	+	0	0	+	+	+
Martini, 2008[520]	RCT	Swine	32ºC	24	PT, aPTT, TEG, ACT	+	+	+	0	0	+	+	+
Mohr, 2013[525]	RCT	Swine	34ºC	40	PT, ROTEM	+	+	+	0	0	0	+	+

Park, 2013[467]	RCT	Rat	33-34ºC	32	ROTEM	+	+	+	0	0	0	+	+
Spronk, 2015[540]	RCT	Swine	32.6 - 33.4 ºC	28	PT, aPTT, ROTEM	+	-	+	0	0	+	+	+

[a] Multivariable adjustment for confounding.

ACT, Activated Clotting Time; AIS, Abbreviated Injury Scale; AMD, Animal model development; BD, Base Deficit; CC, Case-Control; IV, In Vitro; CPR, Cardio pulmonary resuscitation; GSW, Gunshot wound; HS, Haemorrhagic Shock; INR, International Normalised Ratio; ISS, Injury severity score; iTBI, isolated Traumatic Brain Injury; MOI, Mechanism of Injury; NS, Normal Saline; PC, Prospective Cohort; PH, Pre-hospital; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; TA, Trauma Activation; TBV, Total Blood Volume; TI, Tissue Injury; TT, Thrombin Time.

Table 38: Evidence supporting a causal relationship between dilution and coagulopathy

Author, Year	Design	Study Population / Sample	Inclusion Criteria / Exposure	Sample Size	Coagulation measurement	Association	No confounding	**Direct** Temporality	Biological Grad.	Reversibility	**Mechanistic** P. mechanism	**Parallel** Replicability	Similarity
Human Observational Studies													
Aucar, 2003[501]	PC	SC, Civilian, all MOI	Urgent surgical intervention; chest, abdo or extremity	31	Clinical	+	-	+	0	0	0		
Brown, 2013[620]	PC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI ISS>15	1216	INR	+	-	+	0	0	0		
Cohen, 2013[440]	PC	MC, Civilian, all MOI	Blood transfusion within 6hrs	1198	INR, PTT	+	+ [a]	+	+	0	0		
Engels, 2011[502]	RC	SC, Civilian, all MOI	INR>1.3, arrived within 1hr of injury	290	INR	+	+ [a]	+	+	0	0		
Hubetamann, 2011[527]	RC	MC, Civilian, all MOI	German trauma registry, ISS>15, transfused	2702	Quick Test,	+	-	+	0	0	0		
MacLeod, 2014[646]	PC	SC, Civilian, all MOI	TA, admitted for >24hrs	701	PT	+	-	+	0	0	0		
Maegele, 2007[528]	RC	MC, Civilian, all MOI	Entry into German Trauma Registry	8724	Quick Test, Plats	+	-	+	+	0	0		
Neal, 2014[503]	RC	MC, Civilian, Blunt Trauma	Shocked, transfused within 12 hrs, not iTBI	1897	INR, clinical	+	+ [a]	+	+	0	0		
Shaz, 2011[492]	CC	SC, Civilian, all MOI	TA	91	PT	+	-	+	0	0	+		

Sixta, 2012[454]	RC	SC, Civilian, all MOI	Immediate laparotomy	1218	INR	+	+ [a]	+	+	0	0		
Wafaisade, 2010a[566]	RC	MC, Civilian, Blunt trauma	German Trauma Registry, iTBI	3114	PT, Plats	+	-	+	0	0	0		
Wafaisade, 2010b[442]	RC	MC, Civilian, all MOI	German Trauma Registry, ISS >15	1987	PT, Plats	+	-	+	0	0	0		
White, 2015a[626]	CC	SC, Civilian, all MOI	admitted, survived >36hrs	95	INR, aPPT, TEG	+	-	+	0	0	0		
Human Experimental Studies													
Bolliger, 2010[531]	No Control group in - vivo	Patients undergoing coronary artery bypass surgery	Dilution following off-pump coronary artery bypass surgery	9	PT, aPPT, INR, ROTEM	+	+	+	+	0	+	+	+
Caballo, 2013[535]	IV	Healthy whole blood	30% and 60% Dilution	8	PT, aPTT, ROTEM	+	+	+	+	+	+	+	+
Coats, 2006[586]	IV	Healthy whole blood	40% Dilution	12	Sonoclot	+	+	+	0	0	0	+	+
Darlington, 2012[464]	IV	Healthy whole blood	40% Dilution	9	PT, aPTT, TEG	+	+	+	0	+	0	+	+
Schols, 2008[542]	IV	Healthy plasma	20%, 40%, and 60% Dilution	5	TEG	+	+	+	+	+	+	+	+
Animal Experimental Studies													
Chen, 2017[627]	Controlled Trial	Rat	20% Dilution	40	PT, aPTT, ROTEM	+	-	+	0	0	0	+	+
Cho, 2009[628]	AMD	Swine	60% TBV replacement *with NS (3:1)*	37	PT, TEG	+	-	+	0	0	0	+	+
Dickneite, 2009[536]	RCT	Swine	65-70% TBV replacement *with colloid*	47	Bleeding time, PT, aPTT	+	+	+	0	+	+	+	+
Dickneite, 2010[537]	Controlled trial	Swine	65-70% TBV replacement *with colloid*	21	Bleeding time, PT, aPTT	+	+	+	0	+	+	+	+
Fries, 2005[547]	RCT	Swine	65% TBV replacement with colloid (1:1)	14	PT, aPPT, ROTEM	+	+	+	0	+	0	+	+
Fries, 2006[546]	RCT	Swine	65% TBV replacement with colloid (1:1)	20	PT, aPPT, ROTEM	+	+	+	0	+	0	+	+

Grottke, 2010[538]	RCT	Swine	80% TBV replacement with colloid or crystalloid	18	PT, aPTT, ROTEM	+	+	+	0	+	0	+	+
Kheirabadi, 2007[662]	RCT	Rabbit	Crystalloid resuscitation (165ml/min) of uncontrolled haemorrhage	14	PT, aPPT, BT, TEG	+	-	+	0	0	0	+	+
Kiraly, 2006[663]	RCT	Swine	50% TBV replacement with colloid (1:1)	20	PT, aPTT, TEG	+	-	+	0	0	0	+	+
Klemcke, 2005[549]	RCT	Swine	35% TBV replacement with crystalloid (3:1)	54	PT, aPTT, TEG	+	-	+	0	+	0	+	+
Martini, 2006b[635]	RCT	Swine	35% TBV replacement with crystalloid (3:1)	18	TEG	+	-	+	0	0	+	+	+
Martini, 2008[520]	RCT	Swine	60% TBV replacement with colloid (1:1) or crystalloid (3:1)	24	PT, aPTT, TEG, ACT	+	-	+	0	0	+	+	+
Martini, 2013[544]	RCT	Swine	Crystalloid infused at 0.1, 0.25 or 0.75 mL/min	21	TEG	+	-	+	0	0	+	+	+
Nishi, 2013[636]	RCT	Rat	50% TBV replacement with crystalloid (1:1)	18	Sonoclot	+	-	+	+	0	0	+	+
Pawelczyk, 2013[664]	RCT	Rat	50% TBV replacement with colloid (1:1)	25	PT, TEG, Clinical	0	-	0	0	0	0	0	0
Pragst, 2010[548]	RCT	Rabbit	70% TBV replacement with colloid (1:1)	19	PT, BT	+	-	+	0	+	+	+	+
Spronk, 2015[540]	RCT	Swine	30% TBV replacement with crystalloid	28	PT, aPPT, ROTEM	+	-	+	0	+	+	+	+
Torres, 2013[545]	RCT	Rat	40% TBV replacement with crystalloid (3:1) or colloid (1:1)	32	ROTEM	+	-	+	0	0	+	+	+
Watts, 2015[539]	RCT	Swine	30% TBV replacement with crystalloid	24	PT, aPTT, TEG	+	-	+	0	0	0	+	+
Wohlauer, 2012[665]	No Control	Rat	50% Haemodilution with crystalloid (2:1)	10	TEG	0	-	0	0	0	0	0	+

[a] Multivariable adjustment for confounding.

ACT, Activated Clotting Time; aPPT, Activated partial thromboplastin time; AIS, Abbreviated Injury Scale; AMD, Animal model development; BD, Base Deficit; CC, Case-Control; IV, In Vitro; CPR, Cardio pulmonary resuscitation; GSW, Gunshot wound; HS, Haemorrhagic Shock; INR, International Normalised Ratio; ISS, Injury severity score; iTBI, isolated Traumatic Brain Injury; MOI, Mechanism of Injury; NS, Normal Saline; PC, Prospective Cohort; PH, Pre-hospital; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; ROTEM, Rotational thromboelastometry; TA, Trauma Activation; TEG, thromboelastography TBV, Total Blood Volume; TI, Tissue Injury;

Table 39: Evidence supporting a causal relationship between brain injury and coagulopathy

Author, Year	Design	Study Population / Sample	Inclusion Criteria / Exposure	Sample Size	Coagulation measurement			**Direct**			**Mechanistic**	**Parallel**	
						Association	No confounding	Temporality	Biological Grad.	Reversibility	P. mechanism	Replicability	Similarity
Human Observational Studies													
Auer, 1978[554]	Cohort	SC, Civilian	Brain injury admitted to ICU	30	PT, PPT, Plats	+	-	+	0	0	0		
Bredbacka, 1994[555]	Cohort	SC, Civilian, all MOI	iTBI: admitted to neurosurg within 24hrs of injury	20	Quick Score	+	-	+	0	0	0		
Cap, 2011[452]	RC	MC, Military, all MOI	iTBI: head AIS >2, other AIS <3, transfused	1609	INR	+	+ [a]	+	+	0	0		
Carrick, 2005[557]	RC	SC, Civilian, Blunt Trauma	iTBI: GCS <14, other AIS <3	184	PT, PPT, Plats	+	-	+	+	0	0		
Castellino, 2014[576]	CC	SC, Civilian, all MOI	iTBI: head AIS >2, other AIS <3. SBP>90	70	INR, PPT, TEG	+	-	+	0	0	+		
Chhabra, 2013[568]	PR	SC, Civilian, all MOI	iTBI: GCS <13, admitted neurosurgery	208	PT, aPPT	+	+ [a]	+	+	0	0		
Cohen, 2013[440]	PC	MC, Civilian, all MOI	Blood transfusion within 6hrs	1198	INR, PTT	+	-	+	0	0	0		
de Oliveira Manoel, 2015[666]	RC	SC, Civilian, Blunt trauma	TBI = AIS >2, admitted within 6 hours of injury	345	INR, aPPT, Plats, TEG	0	-	0	0	0	0		
Di Battista, 2016[560]	PC	MC, Civilian, All MOI	iTBI: GCS <13, other AIS <3	159	PT, aPPT, Plats	+	-	+	+	0	+		

Genet, 2013[494]	RC	SC, Civilian, all MOI	Full TA	80	INR, aPPT	+	-	+	+	0	0
Goodnight, 1974[553]	PC	SC, Civilian, All MOI	Brain tissue injury, ICU admission <12hrs	26	Quick score, aPPT, TT, Plats	+	-	+	+	0	0
Greuters, 2011[667]	RC	SC, Civilian, All MOI	iTBI: CT confirmed brain injury and other AIS <3	107	INR, aPPT, Plats	+	-	+	0	0	0
Halpern, 2008[559]	RC	MC, Civilian, All MOI	Closed iTBI: head AIS >2, other AIS <3, coagulopathic	542	PT	+	-	+	0	0	0
Hulka, 1996[556]	RC	SC, Civilian, Blunt MOI	Had CT brain, blood sample within 24hrs	159	PT, PPT, Plats	+	-	+	0	0	0
Kearney, 1992[668]	CC	SC, Civilian, All MOI	Head injury, GCS <10, ICU admission	41	PT, PPT, Plats	+	-	+	0	0	0
Keller, 2001[564]	RC	SC, Civilian, Blunt MOI	Age <16, GCS <15 at any time after injury	53	INR, PT, PPT	+	-	+	+	0	0
Lozance, 1998[563]	RC	SC, Civilian, all MOI	iTBI	105	PT, PPT, TT, Plats	+	-	+	+	0	0
Lustenberger, 2010[495]	RC	MC, Civilian, all MOI	iTBI: head AIS >2, other AIS <3, admitted to ICU	132	INR, aPTT, Plats	+	+ [a]	+	+	0	0
MacLeod, 2014[646]	PC	SC, Civilian, all MOI	TA, admitted for >24hrs, brain injury on CT scan	701	PT	+	-	+	+	0	0
May, 1997[647]	RC	SC, Civilian, Blunt Trauma	iTBI: GCS <9, other AIS <3	26	Pt, PPT	+	-	+	+	0	0
Nekludov, 2007[573]	PC	SC, Civilian, Blunt Trauma	iTBI: GCS <9 other AIS <4, admitted to neuroICU	11	INR, bleeding time, Plats	+	-	+	0	0	+
Peiniger, 2012[567]	RC	MC, Civilian, Blunt Trauma	iTBI: head AIS >2, other AIS <3, <14 yrs,	200	Quick value, PTT plat,	+	+ [a]	+	0	0	0
Talving, 2009[441]	PC	SC, Civilian, all MOI	Admitted to ICU, AIS head >2	436	INR, aPPT, Plats	+	+ [a]	+	+	0	0
Turtay, 2010[650]	PC	SC, Civilian, all MOI	Unknown	50	INR, aPPT	+	-	+	+	0	0
van der Sande, 1978[565]	Cohort	SC, Civilian, Blunt Trauma	Unknown	150	Ethanol gelation test	+	-	+	+	0	0
Wafaisade, 2010a[566]	RC	MC, Civilian, Blunt Trauma	iTBI: head AIS >2, other AIS <3	3114	PT, Plats	+	+ [a]	+	0	0	0

Wu, 2014[577]	PC	SC, Civilian, Blunt Trauma	iTBI: Brain injury on CT, extracranial AIS <3, GCS in ED <13	81	INR, aPPT, Plats	+	-	+	+	0	+		
Xu, 2013[456]	RC	SC, Civilian, all MOI	AIS head >2, ISS>15, admitted to ICU <24hrs of injury	223	INR, PPT, aPPT	+	-	+	+	0	0		
Zehtabchi, 2008 [558]	PC	SC, Civilian, all MOI	iTBI: AIS head >2 or intracranial haematoma on CT	276	INR, aPTT	+	-	+	0	0	0		
Human Experimental Studies													
nil													
Animal Experimental Studies													
Castellino, 2014[576]	AMD	Rat	Blunt iTBI	45	TEG	+	+	+	0	0	+	+	+
Donahue, 2014[575]	AMD	Rat	Blunt iTBI	U	PT, aPPT, Plats, TEG	+	+	+	0	+	+	0	0
Tian, 2015[579]	Controlled study	Mouse	Fluid percussion iTBI	U	Clotting time	+	+	+	+	+	+	+	+

[a] Multivariable adjustment for confounding.

ACT, Activated Clotting Time; aPPT, Activated partial thromboplastin time; AIS, Abbreviated Injury Scale; AMD, Animal model development; BD, Base Deficit; CC, Case-Control; IV, In Vitro; CPR, Cardio pulmonary resuscitation; GSW, Gunshot wound; HS, Haemorrhagic Shock; INR, International Normalised Ratio; ISS, Injury severity score; iTBI, isolated Traumatic Brain Injury; MOI, Mechanism of Injury; NS, Normal Saline; PC, Prospective Cohort; PH, Pre-hospital; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; ROTEM, Rotational thromboelastometry; TA, Trauma Activation; TEG, thromboelastography TBV, Total Blood Volume; TI, Tissue Injury; TT, Thrombin Time; U, Unknown.

Table 40: Evidence supporting a causal relationship between fluid type and coagulopathy

Author, Year	Design	Study Population / Sample	Inclusion Criteria / Exposure	Sample Size	Coagulation measurement	Association	No confounding	**Direct** Temporality	Biological Grad.	Reversibility	**Mechanistic** P. mechanism	**Parallel** Replicability	Similarity
Human Observational Studies													
Wafaisade, 2010b[442]	RC	MC, Civilian, all MOI	German trauma registry, ISS>15, colloid to crystalloid ratio	1987	PT, Plats	+	+ [a]	+	0	0	0		
Human Experimental Studies													
Caballo, 2013[535]	IV	Healthy whole blood	0.9% NaCl or RL or Plasma-lyte or HES or 5% Albumin	8	PT, aPTT, ROTEM	+	+	+	+	+	+	+	+
Coats, 2006[586]	IV	Healthy whole blood	0.9% NaCl or Hartmann's, Gelofusin or Haemacel or HES or Albumin	12	Sonoclot	+	+	+	0	0	+	+	+
Darlington, 2012[464]	IV	Healthy whole blood	RL or Hextend	9	PT, aPTT, TEG	+	+	+	0	+	0	+	+
Delano, 2015[596]	RCT	Trauma patients with SBP ≤70 mmHg or SBP 71-90mmHg and HR >108bpm	250 mL of 7.5% NaCl or 7.5% NaCl/6% Dextran 70 or 0.9% NaCl pre-hospital	34	INR, PT, Plats	+	+	+	0	0	+	0	+
Ekseth, 2002[590]	IV	Healthy whole blood	0.9% NaCl or Ringer Acetate or 4% albumin or Dextran 70 or 6% or 10% HES	12	TEG	+	+	+	+	0	0	+	+

Evans, 1996[589]	RCT	Adult trauma patients requiring fluid resuscitation	Haemaccel or RL until fully resuscitated	25	BT, PT, PTT, Plt	+	+	+	0	0	0	0	+
Schols, 2008[542]	IV	Healthy plasma	0.9% NaCl or Gelofusion	5	TEG	+	+	+	+	+	+	+	+
Smith, 2015[599]	RCT	Trauma patients either transfused, intubated or urgent haemorrhage control procedure	0.9% NaCl or Plasma-Lyte A for first 24hrs	18	INR, aPPT, TEG	+	+	+	0	0	0	0	0
Sossdorf, 2009[588]	IV	Healthy whole blood	0.9% NaCl or HES	14	ROTEM	+	+	+	0	0	+	+	+
Tan, 2002[597]	IV	Healthy whole blood	0.9% NaCl or 7.5% NaCl	10	TEG	+	+	+	+	0	0	+	+
Animal Experimental Studies													
Fung, 2012[632]	Controlled trial	Ovine	4% Albumin, NS, PRBCs	14		+	-	+	0	0	0	0	0
Kiraly, 2006[663]	RCT	Swine	0.9% NaCl or LR	20	PT, aPTT, TEG	+	+	+	+	0	+	+	+
Martini, 2013[544]	RCT	Swine	Hextend or LR	21	TEG	+	+	+	0	0	+	+	+
Pawelczyk, 2013[664]	RCT	Rat	LR or whole blood	25	PT, TEG, Clinical	0	+	0	0	0	0	0	0
Torres, 2013[545]	RCT	Rat	LR or Hextend	32	ROTEM	+	+	+	0	0	+	+	+

[a] Multivariable adjustment for confounding.

aPPT, Activated partial thromboplastin time; BT, Bleeding Time; IV, In Vitro; ISS, Injury severity score; LR, Lactated Ringers Solution; 0.9% NaCl, Normal Saline; Plat, Platelets; PPT, Partial Prothrombin time; PRBC, Packed Red Blood Cells; PT, Prothrombin Time; RC, Retrospective Cohort; RCT, Randomised controlled trial; ROTEM, Rotational thromboelastometry; TEG, thromboelastography.