

The BCD Triage Sieve outperforms all existing major incident triage tools

Malik, Nabeela S.; Chernbumroong, Saisakul; Xu, Yuanwei; Vassallo, James; Lee, Justine; Bowley, Douglas M.; Hodgetts, Timothy; Moran, Christopher G; Lord, Janet M; Belli, Antonio; Keene, Damian; Foster, Mark; Gkoutos, Georgios V

DOI:

[10.1016/j.eclinm.2021.100888](https://doi.org/10.1016/j.eclinm.2021.100888)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Malik, NS, Chernbumroong, S, Xu, Y, Vassallo, J, Lee, J, Bowley, DM, Hodgetts, T, Moran, CG, Lord, JM, Belli, A, Keene, D, Foster, M & Gkoutos, GV 2021, 'The BCD Triage Sieve outperforms all existing major incident triage tools: comparative analysis using the UK national trauma registry population', *EClinicalMedicine*, vol. 36, 100888. <https://doi.org/10.1016/j.eclinm.2021.100888>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



Research paper

The BCD Triage Sieve outperforms all existing major incident triage tools: Comparative analysis using the UK national trauma registry population

Nabeela S. Malik^{a,b,c,*}, Saisakul Chernbumroong^{a,d}, Yuanwei Xu^d, James Vassallo^e, Justine Lee^{b,g,h}, Douglas M. Bowley^{b,f,g}, Timothy Hodgettsⁱ, Christopher G Moran^{h,j}, Janet M Lord^{a,b}, Antonio Belli^{a,b,g}, Damian Keene^{f,g}, Mark Foster^{a,e,g}, Georgios V Gkoutos^{a,d,k,l}

^a NIHR Surgical Reconstruction and Microbiological Research Centre (SRMRC), Heritage Building, Queen Elizabeth Hospital, Mindelsohn Way, Edgbaston, Birmingham B15 2TH, UK

^b Institute of Inflammation and Ageing, University of Birmingham, Birmingham B15 2TT, UK

^c 212 (Yorkshire) Field Hospital, Endcliffe Hall, Endcliffe Vale Road, Sheffield S10 3EU, UK

^d Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham B15 2TT, UK

^e Academic Department of Military Emergency Medicine, Royal Centre for Defence Medicine, Mindelsohn Way, Edgbaston, Birmingham B15 2WB, UK

^f Academic Department of Military Surgery & Trauma, Royal Centre for Defence Medicine, Mindelsohn Way, Edgbaston, Birmingham B15 2WB, UK

^g University Hospitals Birmingham, Mindelsohn Way, Edgbaston, Birmingham B15 2WB, UK

^h NHS England London, Skipton House, 80 London Road, London SE1 6LH, UK

ⁱ Army Health, Army Headquarters, Monxton Road, Andover SP11 8HT, UK

^j Nottingham University Hospitals NHS Trust, Derby Road, Nottingham NG7 2UH, UK

^k Institute of Translational Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2TT, UK

^l MRC Health Data Research UK (HDR UK), Midlands Site, B15 2TT UK

ARTICLE INFO

Article History:

Received 22 December 2020

Revised 19 April 2021

Accepted 20 April 2021

Available online 15 May 2021

Keywords:

Major incident

Mass casualty

Prehospital medicine

Triage

Major trauma

Disaster

Injury Severity Score

BCD Triage Sieve

Careflight

START

MSTART

MPTT-24

RAMP

JumpSTART

MIMMS

Life-saving intervention

Military medicine

ABSTRACT

Background: Natural disasters, conflict, and terrorism are major global causes of death and disability. Central to the healthcare response is triage, vital to ensure the right care is provided to the right patient at the right time. The ideal triage tool has high sensitivity for the highest priority (P1) patients with acceptably low over-triage. This study compared the performance of major incident triage tools in predicting P1 casualty status in adults in the prospective UK Trauma Audit and Research Network (TARN) registry.

Methods: TARN patients aged 16+ years (January 2008–December 2017) were included. Ten existing triage tools were applied using patients' first recorded pre-hospital physiology. Patients were subsequently assigned triage categories (P1, P2, P3, Expectant or Dead) based on pre-defined, intervention-based criteria. Tool performance was assessed by comparing tool-predicted and intervention-based priority status.

Findings: 195,709 patients were included; mortality was 7.0% (n=13,601); median Injury Severity Score (ISS) was 9 (IQR 9–17); 97.1% sustained blunt injuries. 22,144 (11.3%) patients fulfilled intervention-based criteria for P1 status, exhibiting higher mortality (12.8% vs. 5.0%, p<0.001), increased intensive care requirement (52.4% vs 5.0%, p<0.001), and more severe injuries (median ISS 21 vs 9, p<0.001) compared with P2 patients. In 16–64 year olds, the highest performing tool was the Battlefield Casualty Drills (BCD) Triage Sieve (Prediction of P1 status: 70.4% sensitivity, over-triage 70.9%, area under the receiver operating curve (AUC) 0.068 [95%CI 0.676–0.684]). The UK National Ambulance Resilience Unit (NARU) Triage Sieve had sensitivity of 44.9%; over-triage 56.4%; AUC 0.666 (95%CI 0.662–0.670). All tools performed poorly amongst the elderly (65+ years).

Interpretation: The BCD Triage Sieve performed best in this nationally representative population; we recommend it supersede the NARU Triage Sieve as the UK primary major incident triage tool. Validated triage category definitions are recommended for appraising future major incidents.

Funding: This study is funded by the National Institute for Health Research (NIHR) Surgical Reconstruction and Microbiology Research Centre. GVG also acknowledges support from the MRC Health Data Research UK (HDRUK/CFC/01). The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care, or the Ministry of Defence.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

* Corresponding author at: NIHR Surgical Reconstruction and Microbiological Research Centre (SRMRC), Heritage Building, Queen Elizabeth Hospital, Mindelsohn Way, Edgbaston, Birmingham, B15 2TH, UK.

E-mail address: nabeelamalik@nhs.net (N.S. Malik).

Research in context

Evidence before this study

Major incidents, ranging from terrorist attacks to large-scale natural disasters, require prioritisation of limited healthcare resources in order to maximise survival amongst those injured. Selection of the optimal triage tool to prioritise patients at the scene of a major incident is an essential component of disaster preparedness and several such tools exist internationally. Those in need of urgent life-saving intervention (Priority 1, P1) are at greatest risk of adverse outcome, hence their timely and accurate identification is of greatest importance, however comparing results from existing studies is limited by lack of consensus on what endpoint best defines P1 status.

Added value of this study

Using 195,000 patients from the UK national trauma registry, this study measured the performance of ten international major incident triage tools, allowing direct comparison in their ability to predict both P1 status and mortality. We validated a system of retrospectively assigning triage categories based on commonly described pre-hospital and hospital interventions. The best performing tool was the UK military's Battlefield Casualty Drills (BCD) Triage Sieve, affording a 24–26% improvement in identifying P1 patients over the current National Ambulance Resilience Unit (NARU) Triage Sieve.

Implications of all the available evidence

Existing consensus-derived definitions of major incident triage categories, which relate directly to healthcare resource utilisation, have been validated. We recommend their use as an endpoint for future evaluations of UK and international major incidents, research and related training. A number of studies have demonstrated that the NARU Triage Sieve used currently by UK ambulance services is not the optimal tool for major incident triage; the BCD Triage Sieve may facilitate a substantial improvement in detecting patients requiring time-critical, life-saving intervention.

1. Introduction

The global incidence of natural disasters, conflict, and terrorism has risen in the last two decades [1,2]; with over two million people dead, many more wounded, and incurring far-reaching economic and societal consequences [1–4]. In the UK, recent major incidents include terrorist-related combined vehicular and stabbing attacks [2], a shrapnel-laden bomb in a concert hall [4], and a large residential fire [2]. Triage, the sorting of casualties according to priority, was conceived during the Napoleonic Wars [5]. Triage enables prioritisation for treatment and onward transfer, and selection of an appropriate destination for definitive medical care [3,6–8]. Those in need of urgent life-saving intervention (Priority 1, or P1, category, also known as Immediate, Red, and Triage category 1 or “T1” internationally) are at greatest risk of potentially preventable adverse outcome, hence their timely and accurate identification is the most important priority of major incident triage [5,8,9]. Accurate triage ensures that limited medical resources are directed towards achieving the greatest possible positive impact for the largest number of people [3,10]. Incorrect triage may fail to identify patients in need of urgent intervention (under-triage); however, its inverse (over-triage) risks overwhelming healthcare facilities with patients who do not require time-critical treatment [3,8–10].

The selection of appropriate major incident triage tools is an important component of disaster and major incident preparedness [1,3,11]. Algorithmic tools used at the scene of a major incident must be quick and simple to apply under challenging circumstances. In the UK, emergency medical services (EMS) currently utilise the National Ambulance and Resilience Unit (NARU) Triage Sieve [12], adapted from the UK military's former MIMMS Triage Sieve [5]. The Battlefield Casualty Drills (BCD) Triage Sieve, used by British soldiers faced with multiple casualties, first appeared in 1998 [13], undergoing serial updates in line with emerging evidence and changes in clinical practice. The most recent update in 2018 (Supplementary Data Fig. 1) incorporated assessment of mental status and a revised respiratory rate threshold (adopted from MPTT-24) [14], a revised heart rate threshold and the rolling of unresponsive patients into the three-quarter prone position. The US-based Simple Triage and Rapid Treatment (START) and modified version MSTART tools [5,15] have been evaluated following several disasters [6,8], and registry-based studies [15,16], demonstrating sensitivity of 85% to 100% in predicting P1

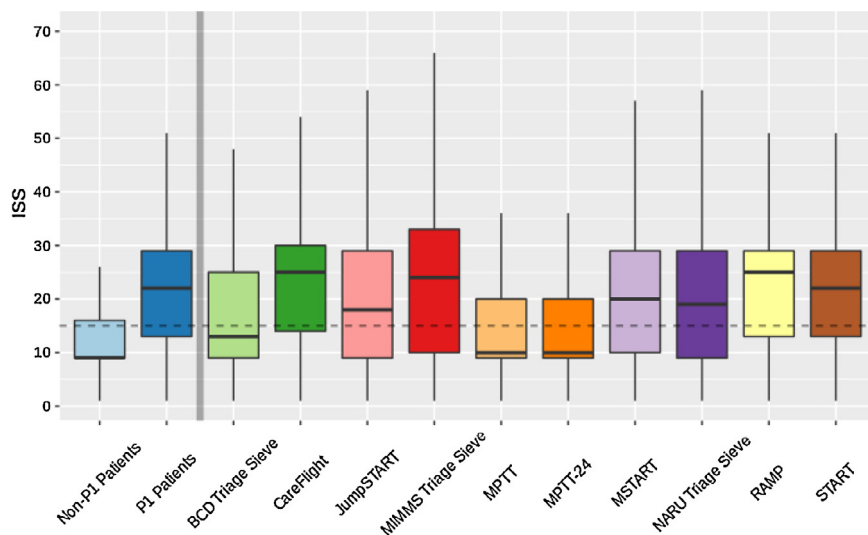


Fig. 1. Distribution of Injury Severity amongst tool-assigned P1 patients (patients aged 16–64 years)

Ledger: ISS=Injury Severity Score. Dotted horizontal line denotes ISS 15. The upper whisker extends from the hinge to the largest value no further than $1.5 \times$ IQR from the hinge; the lower whisker extends from the hinge to the smallest value, at most $1.5 \times$ IQR of the hinge.

status [5]. The Australian CareFlight has demonstrated appreciable sensitivity in predicting P1 status (46–82%) and mortality (AUC 0.852) in events including the 2002 Bali nightclub bombings [5], hospital [15] and trauma registry-based studies [16]. Novel tools include the Modified Physiological Triage Tool (MPTT)[14], MPTT-24 [17], and US-based Rapid Assessment of Mentation and Pulse (RAMP) [18]. JumpSTART, developed for use in children under eight years, has yet to be evaluated in adult patients [5,11]. Few studies examine the performance of triage tools in the elderly [16], who constitute a growing proportion of the UK population [19].

There is a paucity of evidence to guide policy makers in their choice of triage tool [5,8,11]. Conducting prospective studies to directly compare tool performance during major incidents is logistically challenging, given their unpredictable, infrequent nature [2,5,8,10]. Thus, existing evidence comprises post-event evaluations [6,7,10], often limited by small patient numbers with incomplete pre-hospital data [20], simulation studies [8], and studies using hospital [15] or trauma registry patients [14,16] as surrogates for those injured in major incidents. Comparing study results is further limited by lack of consensus on what endpoint best defines P1 status [5,10,15,21]. Some studies have retrospectively utilised ISS > 15 as an endpoint to assign high-acuity status and justify transfer to the uppermost tier of trauma care [10,16]. Whilst ISS and mortality are commonly used, the need for life-saving intervention is recognised as the most appropriate endpoint as this relates directly to resource utilisation in the resource-constrained major incident setting [15,21]. In order to inform major incident triage practice in the UK civilian setting, this study aimed to compare the performance of major incident triage tools applicable at the scene of injury in predicting P1 casualty status amongst adults using the UK trauma registry database. A secondary aim was to assess the utility of a consensus-based system of defining triage categories for the retrospective evaluation of major incident triage [21].

2. Methods

2.1. Overview of study design

This study tests the performance of ten major incident triage tools in predicting P1 status using patients from the UK Trauma Audit and Research Network (TARN) registry. Each triage tool was applied to patients' pre-hospital physiology to determine whether the patient would have been designated P1 status. Patient records were reviewed to determine whether they required time-critical interventions from a pre-defined list, allowing assignment of "actual" triage categories (P1, P2, P3, Expectant or Dead) [21]. Tool performance was assessed by comparing tool-predicted and intervention-based P1 status.

2.2. Selection of participants

Prospectively recorded, anonymised data for TARN registry patients aged 16+ years presenting to hospitals in England and Wales between 1 January 2008 and 31 December 2017 were included. Patients with incomplete pre-hospital physiological data required to apply the triage tools (respiratory rate, heart rate, capillary refill time, Glasgow Coma Score (GCS), and GCS Motor Component) were excluded.

TARN constitutes the largest trauma registry in Europe, receiving data from all UK Major Trauma Centres and Trauma Units [22]. TARN hospital co-ordinators include injured patients fulfilling the following criteria: length of stay over 72 hours, intensive care (ICU) admission and/or in-hospital death [22]. Pre-hospital and in-hospital physiological, demographic, and outcome data are prospectively recorded into a web-based proforma [22]. TARN excludes pre-hospital deaths and elderly patients with isolated femoral neck fractures [22].

2.3. Application of major incident triage tools

It was anticipated that TARN's inclusion criterion of length of stay greater than 72 hours would result in over-representation of elderly patients within the study population. Therefore, to test tool performance, patients were categorised by age into adults (16–64 years) and the elderly (65+ years), consistent with National Health Service configuration. The BCD Triage Sieve (**Supplementary Data Fig. 1**), CareFlight [5], JumpSTART [5], MIMMS Triage Sieve [5], MPTT [14], MPTT-24 [17], MSTART [5], NARU Triage Sieve [12], RAMP [18] and START [5] tools were transcribed into computer code and applied to first recorded pre-hospital physiology to determine whether patients were P1 or non-P1. Tool characteristics including their precise components are summarised in **Table 1**. The US-based SALT [11] and German ASAV [23] tools were considered for inclusion, however these require subjective judgements, limiting reliable retrospective application.

In order to facilitate retrospective application of the triage tools (see **Table 1**), several assumptions were made. By virtue of fulfilling TARN inclusion criteria, all patients were assumed to be non-ambulatory. Patients who had undergone an advanced airway intervention at scene were deemed unable to breathe [24]. A respiratory rate of less than four breaths per minute was regarded undetectable by EMS personnel. The term "catastrophic haemorrhage" utilised by the BCD and NARU Triage Sieve, and MPTT-24 could not be applied retrospectively as this field is not captured by TARN. Patients with a systolic blood pressure of 90 mmHg and over were regarded as having a palpable radial pulse [25]. Patients were deemed unconscious if their GCS score was less than or equal to eight and those with a GCS of less than 12 were deemed unresponsive to voice [26]. Ability to follow commands was equated to GCS Motor Score of six by convention. In applying JumpSTART, a GCS Motor Score of three or less was regarded as equivalent to "inappropriate response to painful stimulus (*e.g.* positioning) or unresponsive to noxious stimulus [26]."

2.4. Outcome measures

The primary outcome measure was the ability of triage tools to predict P1 status, defined as the need for time-critical lifesaving intervention(s) [21]. Each patient was assigned a triage category (Dead, Expectant, P1, P2 or P3) based on a pre-defined system utilising EMS and hospital-based interventions described by Lerner *et al.*, using equivalent TARN terminology (see **Supplementary Data Table 1**) [21]. As TARN does not include patients with chemical, biological, radiological, and nuclear injuries, criteria relevant to this injury mechanism were not included [21,22]. TARN records the timing of hospital arrival and each intervention, allowing incorporation of this interval into the time-critical definitions constituting P1 status. To assess the validity of Lerner's system of classification [21], patients within each category were compared with regard to mortality, ICU admission, hospital length of stay (LOS) and ISS.

Secondary outcome measures included prediction of mortality and ISS > 15, and distribution of ISS amongst tool-assigned P1 patients, which may provide further discriminative value and appreciation of tool characteristics.

2.5. Data processing and analyses

TARN data were received in SPSS Version 24.0 (Armonk NY: IBM Corp 2015) and processed using R software (Version 3.6, R Core Team, New Zealand, 2000). Non-parametric data are presented as median and interquartile range; categorical data as frequency and percent. To indicate whether the differences between P1 and P2 patients as designated by Lerner's criteria [21] were statistically significant, the Chi-squared test (comparing mortality and ICU admission) and Mood's median test [27] (comparing ISS) were utilised. Performance characteristics included sensitivity, specificity, positive

Table 1

Summary of triage tool characteristics.

Tool	Description and geographical use	Tool components 1st step	2nd step	3rd step	4th step	5th step	6th step	7th step	Interventions permitted
Battlefield Casualty Drills (BCD) Triage Sieve	Current UK military tool for use in adults (introduced in 1998, revised 2018).	Catastrophic haemorrhage?	Walking?	Breathing?	Responds to voice?	Breathing rate between 12–23	Heart Rate more than 100	–	Apply tourniquet, open airway, place casualty in the $\frac{3}{4}$ prone recovery position
CareFlight	Australian tool used in adults and children (introduced in 2001).	Walks?	Obeys command?	Palpable radial pulse? OR Breathes with open airway?	–	–	–	–	Open airway
Jump Simple Triage and Rapid Treatment (JumpSTART)	United States, used in several states in children (introduced in 2001).	Able to walk?	Spontaneous breathing (check radial pulse if apnoeic)	Respiratory rate <15 or >45	Palpable pulse?	Neurological Assessment (AVPU)	–	–	Airway positioning, 5 rescue breaths if apnoeic
Major Incident Medical Management and Support (MIMMS) Triage Sieve	Former UK military adult triage tool (introduced in 1995).	Walking	Breathing	Respiratory rate <10 or ≥ 30	Capillary refill >2 seconds	–	–	–	Open airway
Modified Physiological Triage Tool (MPTT)	UK-based tool* modelled in a military cohort (described in 2017).	Walking?	Breathing?	Respiratory rate <12 or ≥ 22	Heart rate ≥ 100	GCS <14	–	–	–
Modified Physiological Triage Tool 24 (MPTT-24)	UK-based tool*, modification of MPTT (described in 2017).	Catastrophic Haemorrhage?	Walking?	Breathing?	Responds to voice	Respiratory rate <12 or ≥ 24	Heart rate ≥ 100	–	Apply tourniquet or haemostatic dressing
Modified Simple Triage and Rapid Treatment (MSTART)	United States, modification of START (described in 2006).	Able to walk?	Spontaneous breathing	Respiratory rate >30	Radial pulse absent	Obeys commands	–	–	Position airway
National Ambulance and Resilience Unit (NARU) Triage Sieve	Current UK civilian adult tool, adapted from the MIMMS Triage Sieve (this version was introduced in 2013)	Catastrophic haemorrhage	Are they injured	Walking	Breathing	Unconscious	Respiratory rate <10 or ≥ 30	Pulse >120 or capillary refill >2 sec	Apply tourniquet/ haemostatic dressing, open airway, place in recovery position
Rapid Assessment of Mentation and Pulse (RAMP)	United States, used by the Rocky Mountain Fire Department, Colorado (introduced in 2018).	Casualty without signs of obvious death	Casualty follows commands	Radial pulse present?	–	–	–	–	Control massive haemorrhage, open airway, chest decompression
Simple Triage and Rapid Treatment (START)	United States (introduced in 1983).	Able to walk?	Spontaneous breathing	Respiratory rate >30	Capillary refill >2 sec	Obeys commands	–	–	Position airway
Sort, Assess, Life-saving interventions, Treatment/ Transport (SALT)	United States (introduced in 2008 by the Centre of Disease Control).	Sort**	Breathing	Obeys commands or makes purposeful movements?	Has peripheral pulse?	Not in respiratory distress?	Major haemorrhage is controlled?	Likely to survive given current resources?	Control major haemorrhage, open airway (if child, consider 2 rescue breaths), chest

(continued on next page)

Table 1 (Continued)

Tool	Description and geographical use	Tool components	1st step	2nd step	3rd step	4th step	5th step	6th step	7th step	Interventions permitted
Amberg-Schwandt Algorithm*** (ASAV)	German adaptation of the MSTART (described in 2013).	Ambulating?	Deadly injured?	Breathing difficulties?	Spurting haemorrhage?	Radial pulse absent?	Unable to follow simple commands?			decompression, auto injector antidotes Keep airway open

Ledger: Respiratory rate and heart rate are measured as breaths and beats per minute, respectively. AVPU refers to the Alert, Voice, Pain, Unresponsive scale; GCS=Glasgow Coma Score. All tools described are applicable at the scene of a major incident (primary triage tools). *Has yet to undergo practical use or implementation studies. SALT and ASAV were not evaluated in this study as there were major limitations in applying these retrospectively. **SALT involves sorting according to the following: walk, wave/purposeful movement, still/obvious life threat; as well as the subjective judgements: "Minor injuries only?" and "Likely to survive given current resources?" ***ASAV includes the subjective judgement "Deadly injured?" and assessment of breathing status as follows: "airway obstructed, bradypnoea, tachypnoea, apnoea, dyspnoea, cyanosis."

Table 2
Patient and injury characteristics (n=195,709).

Characteristic	n (%)
Gender	
Male	104,019 (53.1%)
Female	91,690 (46.9%)
Missing data	0 (0.0%)
Injury Severity Score (ISS)	
Median (IQR)	9 (9–17)
Missing data	0 (0.0%)
Age	
Median (IQR), years	66.2 (47.3–83.0)
Patients aged 16–64 years	95,306 (48.7%)
Patients aged 65+ years	100,403 (51.3%)
Missing data	0 (0.0%)
Discharge status	
Alive	182,107 (93.0%)
Dead	13,601 (7.0%)
Missing data	1 (0.0%)
Mode of injury	
Blunt	190,048 (97.1%)
Penetrating	5660 (2.9%)
Missing data	1 (0.0%)
Mechanism of injury	
Fall less than 2m	113,319 (57.9%)
Vehicle Incident/Collision	41,590 (21.3%)
Fall more than 2m	25,194 (12.9%)
Blow(s)	6827 (3.5%)
Stabbing	4105 (2.1%)
Other	2609 (1.3%)
Crush	1355 (0.7%)
Shooting	440 (0.2%)
Blast	142 (0.1%)
Burn	128 (0.1%)
Missing data	0 (0.0%)

predictive value, negative predictive value, under-triage (1-sensitivity), over-triage (1-positive predictive value), and Area Under the Receiver Operating Curve (AUC). 95% confidence intervals were calculated using the Wilson Score with continuity correction for binomial proportions, and DeLong's Algorithm for comparing AUC curves [28]. To estimate bias, patients included in the study were compared to those excluded with respect to clinical and demographic characteristics (Supplementary Data Table 2). A value of $p < 0.05$ was considered statistically significant.

Ethical approval: The UK Health Research Authority Patient Information Advisory Group (Section 20) have granted ethical approval and waived the requirement for individual participant consent for research using anonymised TARN data.

Role of Funding: The funding source had no role to play in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

3. Results

3.1. Characteristics of the study population

TARN captured 215,632 patients. 19,923 (9.2%) patients were excluded due to incomplete data; hence, 195,709 patients were included. Patient and injury characteristics are summarised in Table 2. There was a slight male preponderance (53.1% male vs. 46.9% female). Elderly patients constituted approximately half of the study population (n=100,403, 51.3%). Mortality was 7.0% (n=13,601), median ISS was 9 (IQR 9–17). Blunt-injury patients accounted for 97.1% of patients, with low-level falls (n=113,319, 57.9%) and vehicle collision (n=41,590, 21.3%) recorded as the most prevalent injury mechanisms. Penetrating trauma constituted only 2.9% (n=5660), largely comprising stabbings (n=4105, 2.1%). Comparison between included and excluded patients is shown in Supplementary Data Table 2.

Table 3

Comparison of outcome characteristics between patients in each triage category.

Triage category	Total, n (%)	Mortality, n (%)	Intensive care admission, n (%)	Length of stay (days), median (IQR)	ISS, median (IQR)
Dead	282 (0.1)	282 (100.0)	145 (51.4)	1 [1, 3]	27 [25, 41]
Expectant	1879 (1.0)	1862 (99.1)	1021 (54.3)	1 [1, 4]	26 [25, 38]
Priority 1 (Immediate)	22,144 (11.3)	2839 (12.8)	11,593 (52.4)	11 [4, 26]	21 [10, 29]
Priority 2 (Urgent)	171,404 (87.6)	8618 (5.0)	8661 (5.0)	10 [5, 18]	9 [9, 16]

Ledges: IQR=interquartile range, ISS=Injury Severity Score.

Table 4

Breakdown of time-critical life-saving interventions constituting Priority 1 status.

Subcomponents of the Priority 1 triage category	n (% of P1 patients)
An advanced airway intervention (e.g. intubation, LMA, surgical airway) performed in the pre-hospital setting or within 4 hours of arrival at hospital	18,890 (85.3%)
Chest tube placed within 2 hours of arrival at hospital	4301 (19.4%)
Neurological, vascular, or haemorrhage-controlling surgery to the head, neck or torso performed within 4 hours of arrival to hospital	3427 (15.5%)
Arrived in the ED with uncontrolled haemorrhage	1979 (8.9%)
Limb-conserving surgery performed within 4 hours of arrival at hospital on a limb that was found to be pulseless distal to the injury prior to surgery	606 (2.7%)
IV vasopressors administered within 2 hours of arrival at hospital	361 (1.6%)
Patient who required EMS initiation of CPR (i.e. had a cardiac arrest) during transport, in the ED, or within 4 hours of arrival at a hospital	182 (0.8%)
Escharotomy performed on a patient with burns within 2 hours of arrival at a hospital	3 (0.0%)
Total number of P1 patients	22,144 (100.0%)

Ledges: There is overlap between life-saving interventions (LSI): 78.0% (n=17,272) of P1 patients required one LSI, 15.9% (n=3520) required two LSI, and 6.0% (n=1328) required 3 or more LSI.

3.2. Designation of intervention-based triage categories (using Lerner's criteria)

A minority of patients met criteria for the "Dead" category (n=282, 0.1%) whilst 1.0% (n=1879) of the study population were classed as Expectant due to non-survivable burns (n=36) and catastrophic head injury (n=1843) (**Table 3**). 22,144 (11.3%) of patients satisfied criteria for P1 status; most commonly requiring airway intervention (n=18,890, 85.3%), emergent chest tube placement (n=4301, 19.4%), and emergency surgery (n=3427, 15.5%) (**Table 4**). No TARN patients met criteria for the minimally injured P3 group (**Supplementary Data Table 1**). The remaining patients (n=171,402, 87.0%) were assigned P2, forming the largest triage category.

Comparison of patient characteristics across the four triage categories revealed important differences (**Table 3**). Mortality was universal in those classed "Dead" and 99.1% in those labelled Expectant, both groups exhibited very severe injuries (median ISS 27 and 26, respectively). P1 patients demonstrated more than double the mortality (12.8% vs. 5.0%; p<0.001), ten times the rate of ICU admission (52.4% vs. 5.0%; p<0.001), and more severe injuries (median ISS 21 vs. median ISS 9; p<0.001, respectively) compared to patients designated P2.

Table 5

Tool Performance in Predicting Intervention-based Priority 1 Status in Adults aged 16–64 years.

Tool	Sensitivity	Specificity	PPV	NPV	Undertriage (1-sensitivity)	Overtriage (1-PPV)	AUC
BCD Triage Sieve	70.4 [69.7, 71.1]	65.6 [65.3, 66.0]	29.1 [28.6, 29.6]	91.7 [91.5, 91.9]	29.6 [28.9, 30.3]	70.9 [70.4, 71.4]	0.680 [0.676, 0.684]
CareFlight	43.3 [42.6, 44.1]	92.8 [92.7, 93.0]	54.8 [53.9, 55.7]	89.1 [88.9, 89.3]	56.7 [55.9, 57.4]	45.2 [44.3, 46.1]	0.681 [0.677, 0.685]
JumpSTART	46.8 [46.1, 47.6]	89.3 [89.0, 89.5]	46.6 [45.8, 47.4]	89.3 [89.1, 89.6]	53.2 [52.4, 53.9]	53.4 [52.6, 54.2]	0.681 [0.676, 0.685]
MIMMS Triage Sieve	41.8 [41.0, 42.5]	93.4 [93.3, 93.6]	56.0 [55.1, 56.9]	88.9 [88.7, 89.1]	58.2 [57.5, 59.0]	44.0 [43.1, 44.9]	0.676 [0.672, 0.680]
MPIT	49.9 [49.1, 50.7]	59.1 [58.7, 59.4]	19.6 [19.2, 20.0]	85.5 [85.2, 85.8]	50.1 [49.3, 50.9]	80.4 [80.0, 80.8]	0.545 [0.541, 0.549]
MPIT-24	47.9 [47.1, 48.7]	62.9 [62.6, 63.2]	20.6 [20.1, 21.0]	85.8 [85.5, 86.1]	52.1 [51.3, 52.9]	79.4 [79.0, 79.9]	0.554 [0.550, 0.558]
MSTART	57.2 [56.5, 58.0]	89.0 [88.8, 89.3]	51.1 [50.4, 51.9]	91.2 [91.0, 91.4]	42.8 [42.0, 43.5]	48.9 [48.1, 49.6]	0.731 [0.727, 0.735]
NARU Triage Sieve	44.9 [44.1, 45.7]	88.4 [88.2, 88.6]	43.6 [42.9, 44.4]	88.9 [88.7, 89.1]	55.1 [54.3, 55.9]	56.4 [55.6, 57.1]	0.666 [0.662, 0.670]
RAMP	39.4 [38.6, 40.1]	93.3 [93.1, 93.5]	54.1 [53.2, 55.0]	88.5 [88.3, 88.7]	60.6 [59.9, 61.4]	45.9 [45.0, 46.8]	0.663 [0.660, 0.667]
START	53.7 [52.9, 54.5]	90.9 [90.7, 91.1]	54.2 [53.4, 55.0]	90.7 [90.5, 90.9]	46.3 [45.5, 47.1]	45.8 [45.0, 46.6]	0.723 [0.719, 0.727]

Ledges: Results are accompanied by 95% confidence intervals. PPV=positive predictive value, NPV=negative predictive value. AUC=Area Under the Receiver Operating Curve.

Table 6

Tool Performance in Predicting Intervention-based Priority 1 Status amongst the Elderly (aged 65± years).

Tool	Sensitivity	Specificity	PPV	NPV	Undertriage (1-sensitivity)	Overtriage (1-PPV)	AUC
BCD Triage Sieve	56.7 [55.5, 57.9]	72.7 [72.4, 73]	12.1 [11.7, 12.5]	96.2 [96.1, 96.3]	43.3 [42.1, 44.5]	87.9 [87.5, 88.3]	0.647 [0.641, 0.653]
CareFlight	33.5 [32.3, 34.7]	93.4 [93.3, 93.6]	25.3 [24.4, 26.3]	95.5 [95.4, 95.6]	66.5 [65.3, 67.7]	74.7 [73.7, 75.6]	0.635 [0.629, 0.641]
JumpSTART	36.1 [34.9, 37.3]	90.7 [90.5, 90.9]	20.5 [19.7, 21.2]	95.5 [95.4, 95.7]	63.9 [62.7, 65.1]	79.5 [78.8, 80.3]	0.634 [0.628, 0.640]
MIMMS Triage Sieve	34.7 [33.5, 35.9]	92.8 [92.7, 93.0]	24.3 [23.4, 25.2]	95.5 [95.4, 95.7]	65.3 [64.1, 66.5]	75.7 [74.8, 76.6]	0.638 [0.632, 0.644]
MPIT	45.4 [44.1, 46.6]	66.4 [66.1, 66.7]	8.2 [7.9, 8.5]	94.8 [94.7, 95.0]	54.6 [53.4, 55.9]	91.8 [91.5, 92.1]	0.559 [0.553, 0.565]
MPIT-24	43.1 [41.9, 44.3]	69.9 [69.6, 70.2]	8.7 [8.4, 9.0]	94.9 [94.7, 95.0]	56.9 [55.7, 58.1]	91.3 [91.0, 91.6]	0.565 [0.559, 0.571]
MSTART	48.6 [47.4, 49.9]	88.5 [88.3, 88.7]	21.8 [21.2, 22.5]	96.3 [96.2, 96.4]	51.4 [50.1, 52.6]	78.2 [77.5, 78.8]	0.686 [0.679, 0.692]
NARU Triage Sieve	33.2 [32.1, 34.4]	89.6 [89.4, 89.8]	17.5 [16.9, 18.2]	95.3 [95.2, 95.4]	66.8 [65.6, 67.9]	82.5 [81.8, 83.1]	0.614 [0.609, 0.620]
RAMP	31.3 [30.1, 32.4]	93.7 [93.5, 93.9]	24.7 [23.8, 25.7]	95.4 [95.2, 95.5]	68.7 [67.6, 69.9]	75.3 [74.3, 76.2]	0.625 [0.619, 0.631]
START	45.9 [44.7, 47.2]	89.9 [89.7, 90.1]	23.2 [22.5, 24.0]	96.2 [96.0, 96.3]	54.1 [52.8, 55.3]	76.8 [76.0, 77.5]	0.679 [0.673, 0.686]

Ledges: Results are accompanied by 95% confidence intervals. PPV=positive predictive value, NPV=negative predictive value. AUC=Area Under the Receiver Operating Curve.

Table 7

Tool performance in predicting mortality in adults aged 16–64 years.

Tool	Sensitivity	Specificity	PPV	NPV	Undertriage (1-sensitivity)	Overtriage (1-PPV)	AUC
BCD Triage Sieve	85.2 [83.8, 86.6]	60.9 [60.5, 61.2]	5.6 [5.4, 5.8]	99.3 [99.3, 99.4]	14.8 [13.4, 16.2]	94.4 [94.2, 94.6]	0.730 [0.723, 0.738]
CareFlight	69.6 [67.8, 71.4]	88.3 [88.1, 88.6]	14.0 [13.4, 14.6]	99.1 [99.0, 99.1]	30.4 [28.6, 32.2]	86.0 [85.4, 86.6]	0.790 [0.781, 0.799]
JumpSTART	70.0 [68.2, 71.8]	84.7 [84.5, 84.9]	11.1 [10.6, 11.6]	99.0 [99.0, 99.1]	30.0 [28.2, 31.8]	88.9 [88.4, 89.4]	0.774 [0.765, 0.783]
MIMMS Triage Sieve	63.3 [61.4, 65.2]	88.9 [88.7, 89.1]	13.5 [12.9, 14.2]	98.9 [98.8, 99.0]	36.7 [34.8, 38.6]	86.5 [85.8, 87.1]	0.761 [0.752, 0.771]
MPTT	34.2 [32.3, 36.1]	57.4 [57.0, 57.7]	2.1 [2.0, 2.3]	97.0 [96.8, 97.1]	65.8 [63.9, 67.7]	97.9 [97.7, 98.0]	0.458 [0.448, 0.467]
MPTT-24	33.4 [31.6, 35.3]	61.0 [60.6, 61.3]	2.3 [2.1, 2.4]	97.1 [97.0, 97.2]	66.6 [64.7, 68.4]	97.7 [97.6, 97.9]	0.472 [0.463, 0.481]
MSTART	77.3 [75.6, 78.9]	82.9 [82.7, 83.2]	11.0 [10.5, 11.5]	99.3 [99.2, 99.3]	22.7 [21.1, 24.4]	89.0 [88.5, 89.5]	0.801 [0.793, 0.809]
NARU Triage Sieve	72.7 [70.9, 74.4]	84.3 [84.1, 84.6]	11.2 [10.8, 11.7]	99.1 [99.1, 99.2]	27.3 [25.6, 29.1]	88.8 [88.3, 89.2]	0.785 [0.776, 0.794]
RAMP	50.6 [48.6, 52.6]	88.9 [88.7, 89.1]	11.1 [10.5, 11.7]	98.5 [98.4, 98.6]	49.4 [47.4, 51.4]	88.9 [88.3, 89.5]	0.698 [0.688, 0.707]
START	75.3 [73.6, 77.0]	85.1 [84.8, 85.3]	12.1 [11.6, 12.6]	99.2 [99.1, 99.3]	24.7 [23.0, 26.4]	87.9 [87.4, 88.4]	0.802 [0.794, 0.810]

Ledger: Results are accompanied by 95% confidence intervals. PPV=positive predictive value, NPV=negative predictive value. AUC=Area Under the Receiver Operating Curve.

Table 8

Tool Performance in Predicting Mortality amongst the Elderly (aged 65± years).

Tool	Sensitivity	Specificity	PPV	NPV	Undertriage (1-sensitivity)	Overtriage (1-PPV)	AUC
BCD Triage Sieve	49.7 [48.8, 50.7]	73.4 [73.1, 73.7]	18.8 [18.4, 19.3]	92.2 [92.0, 92.4]	50.3 [49.3, 51.2]	81.2 [80.7, 81.6]	0.616 [0.611, 0.621]
CareFlight	26.9 [26.1, 27.8]	94.1 [93.9, 94.2]	36.1 [35.0, 37.1]	91.2 [91.0, 91.4]	73.1 [72.2, 73.9]	63.9 [62.9, 65.0]	0.605 [0.601, 0.609]
JumpSTART	23.5 [22.7, 24.3]	90.6 [90.4, 90.8]	23.6 [22.9, 24.5]	90.5 [90.3, 90.7]	76.5 [75.7, 77.3]	76.4 [75.5, 77.1]	0.571 [0.566, 0.575]
MIMMS Triage Sieve	21.5 [20.8, 22.3]	92.7 [92.5, 92.9]	26.8 [25.8, 27.7]	90.5 [90.3, 90.7]	78.5 [77.7, 79.2]	73.2 [72.3, 74.2]	0.571 [0.567, 0.575]
MPTT	48.6 [47.7, 49.5]	67.5 [67.2, 67.8]	15.6 [15.2, 16.0]	91.4 [91.2, 91.6]	51.4 [50.5, 52.3]	84.4 [84.0, 84.8]	0.580 [0.575, 0.585]
MPTT-24	45.9 [44.9, 46.8]	71.0 [70.7, 71.3]	16.4 [16.0, 16.8]	91.4 [91.2, 91.6]	54.1 [53.2, 55.1]	83.6 [83.2, 84.0]	0.584 [0.579, 0.589]
MSTART	35.7 [34.8, 36.6]	88.9 [88.7, 89.1]	28.4 [27.7, 29.2]	91.8 [91.6, 92.0]	64.3 [63.4, 65.2]	71.6 [70.8, 72.3]	0.623 [0.618, 0.627]
NARU Triage Sieve	29.4 [28.6, 30.3]	90.4 [90.2, 90.6]	27.5 [26.7, 28.4]	91.2 [91.0, 91.4]	70.6 [69.7, 71.4]	72.5 [71.6, 73.3]	0.599 [0.595, 0.604]
RAMP	25.0 [24.2, 25.9]	94.3 [94.1, 94.4]	35.1 [34.1, 36.2]	91.0 [90.8, 91.2]	75.0 [74.1, 75.8]	64.9 [63.8, 65.9]	0.597 [0.592, 0.601]
START	33.5 [32.7, 34.4]	90.3 [90.1, 90.5]	30.1 [29.3, 30.9]	91.6 [91.5, 91.8]	66.5 [65.6, 67.3]	69.9 [69.1, 70.7]	0.619 [0.615, 0.624]

Ledger: Results are accompanied by 95% confidence intervals. PPV=positive predictive value, NPV=negative predictive value. AUC=Area Under the Receiver Operating Curve.

3.3. Triage tool performance

3.3.1. Prediction of P1 status

Tool performance in predicting P1 status in adults aged 16–64 years is shown in **Table 5**. The NARU Triage Sieve exhibited sensitivity of 44.9% in predicting P1 status with associated over-triage of 56.4% and an AUC of 0.666 (95% CI 0.662–0.670). The BCD Triage Sieve demonstrated the highest sensitivity (70.4%) in predicting P1 status with associated over-triage of 70.9% and an AUC of 0.680 (95% CI 0.676–0.684).

MSTART demonstrated the next highest sensitivity in predicting P1 status (57.2%) with the most favourable AUC of 0.731 (95% CI 0.727–0.735); followed closely by parent tool START. The novel RAMP demonstrates lower overall performance to CareFlight. MPTT and MPTT-24 demonstrated moderate sensitivity (49.9% and 47.9%), the highest over-triage rates (80.4% and 79.4%) and lowest specificity (59.1% and 62.9%). Amongst adults aged 65+ years, tools demonstrated high over-triage and all tools except for the BCD Triage Sieve achieved less than 50% sensitivity in predicting P1 status (**Table 6**).

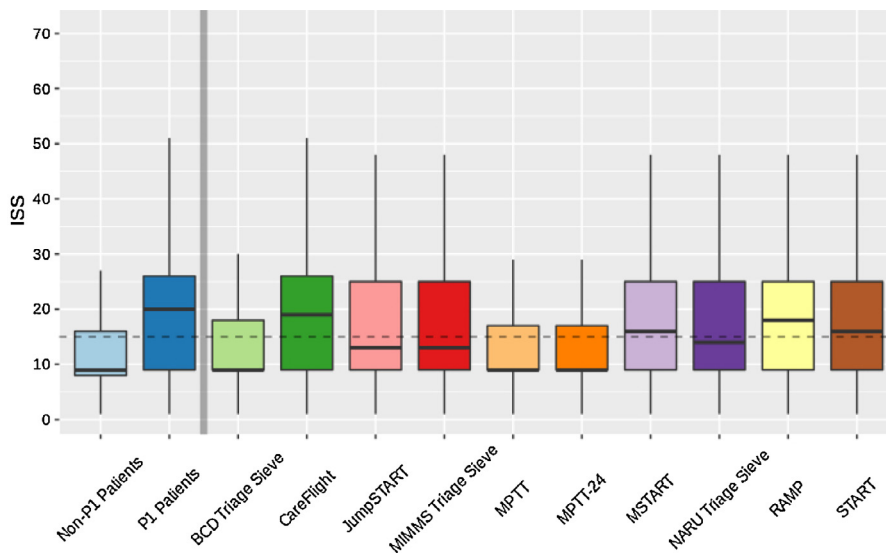


Fig. 2. Distribution of Injury Severity amongst tool-assigned P1 patients (patients aged 65± years)

Ledger: ISS=Injury Severity Score. Dotted horizontal line denotes ISS 15. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge; the lower whisker extends from the hinge to the smallest value, at most 1.5 * IQR of the hinge.

3.3.2. Prediction of mortality

Amongst 16–64 year olds, there was greater variation in tool performance in predicting mortality (Table 7). The BCD Triage Sieve exhibited the highest sensitivity (85.2%) in predicting mortality, followed by MSTART (77.3%), START (75.3%), and NARU Triage Sieve (72.7%); all four tools attained comparable AUC (0.730–0.802). MPTT and MPTT-24 predict mortality with sensitivity of 34.2% and 33.4%, respectively, with AUC below 0.5. All tools performed poorly (sensitivity 21.5–49.7%; AUC 0.571–0.623) in predicting mortality in the elderly (Table 8).

3.3.3. Injury Severity Score

Patients with ISS > 15 totalled 62,402; of these, only 24.1% (n=15,058) met criteria for intervention-based P1 status whilst 75.9% (n=47,344) were non-P1. One third (n=7086) of intervention-based P1 patients had an ISS ≤ 15. There was great variation in the distribution of ISS amongst tool-assigned P1 patients (Figs. 1 and 2). Tool performance in predicting ISS > 15 is included in Supplementary data Tables 3 and 4.

4. Discussion

Globally, natural disasters, conflict, and terrorism pose significant and often unexpected threats, incurring substantial societal and economic impact [1,2,29]. Review of the 2017 Manchester Arena attack highlighted the importance of well-co-ordinated, multiagency collaboration in processing casualties and the negative impacts of an inadequate EMS response [4]. Meticulous disaster planning, including selection of an effective triage tool, is crucial to maximising survival [2,3,8,10,29]. This study measured the performance of ten major incident triage tools in predicting P1 status and mortality using a nationally representative UK adult patient population. A system for defining major incident triage categories relating directly to healthcare resource utilisation has been validated, yielding patient groups with distinct clinical characteristics. ISS > 15 correlated poorly with the need for life-saving intervention, making this a suboptimal endpoint for evaluating major incidents. The NARU Triage Sieve, currently used in UK practice, is outperformed by several other triage tools. The military-derived BCD Triage Sieve demonstrates the greatest sensitivity in predicting both P1 status (sensitivity 70.4%, AUC 0.680, over-triage 70.9%) and mortality (sensitivity 85.2%, AUC 0.730). All tools performed poorly amongst the elderly.

Whilst the American College of Surgeons has established standards for the performance of pre-hospital triage tools for individual patients (acceptable rates of under-triage of up to 5% and 25–50% over-triage, respectively) [7], no national or international standards exist to govern tool performance in major incidents [3,11,12]. From a clinical perspective, under-triage of critically unwell (P1) patients leads to absolute harm arising from delayed care or transfer to an inappropriate medical facility. Conveying critically injured patients to the highest tier of trauma care within trauma networks is associated with decreased mortality: in the UK, bypass of local hospitals to designated Major Trauma Centres has been associated with a 19% increase in the adjusted odds of survival following severe injury [30]. However, minimising over-triage in the resource-constrained major incident setting is also crucially important as overwhelming medical facilities with patients suffering non-critical injury can impair care for those requiring time-critical interventions [8–10]: a study encompassing 3357 casualties from 220 bombing incidents demonstrated a direct linear relationship between over-triage and critical mortality [9]. Policymakers are therefore likely to favour the clinically relevant measures of under-triage and over-triage over AUC, which offers an aggregate measure of performance, selecting tools that align with local casualty distribution plans and available resources [3,4]. Whilst the need for life-saving intervention is of prime importance in guiding resource allocation in major incidents, predicting all-cause in-

hospital mortality in patients with diverse, multi-system injuries is an additional useful measure of tool performance, given that the priority is to maximise overall survival [3,5,12,16]. The NARU Triage Sieve currently used in UK practice performs suboptimally in predicting both P1 status and mortality. Based on the markedly superior sensitivity in predicting P1 status and mortality in patients in this study, we recommend the BCD Triage Sieve as the primary triage tool for adult patients in the UK major incident setting. This change may afford a 24–26% improvement in detecting patients at the scene of a major incident who require time-critical life-saving intervention (and an additional 13–20% of patients who may suffer in-hospital mortality), facilitating their immediate transfer to the highest tier of trauma care and thereby maximising their chances of survival. The BCD Triage Sieve is similar in format to the NARU Triage Sieve, a potential advantage in the retraining of EMS personnel. Furthermore, omission of the need to measure capillary refill may reduce the time taken to perform triage.

The MSTART, START, and CareFlight tools demonstrate considerable sensitivity in predicting mortality, consistent with the findings of a large US registry-based study [16]. CareFlight involves four simple qualitative assessments, apparently achievable within 15 seconds [5]. CareFlight's previously demonstrated superior performance in children confers the potential advantage that a single tool used across all ages would simplify EMS training and practice [16]. Such strengths render it a viable choice for non-clinical UK emergency services personnel, where its relative ease of applicability and low over-triage rate may prove advantageous. In adults, JumpSTART is substantially outperformed by several other tools, limiting its potential utility as a single tool for use across all ages. MPTT and MPTT-24 have comparable over-triage rates to the BCD Sieve, however the BCD Sieve offers a 20.5–22.5% sensitivity advantage in predicting the need for life-saving intervention. Amongst the elderly, all tools performed poorly, consistent with the findings of a large US registry-based study [16]; this may be attributable to age-related physiological changes, chronic illness and polypharmacy. The elderly represent 18.3% of the total UK population [19], yet constitute 51.3% of the study population (51.3%), likely due to the TARN inclusion criterion of admission exceeding 72 hours²²: this over-representation has been mitigated by analysing this cohort as a separate subgroup. Further research is needed to determine how triage can be improved in the elderly.

A key strength of this study is the use of a nationally representative sample of patients, with detailed physiological data, injured by a range of mechanisms. Our findings are therefore applicable to UK adults involved in all-hazard major incidents. ISS has again been shown to correlate poorly with resource use and need for intervention [31]. Our study has externally validated a consensus-derived definition of triage categories [21]. Application of Lerner's criteria defined P1 patients with characteristics (mortality 12.8%, 52.4% ICU admission, median ISS of 21) appropriate for the highest tier of trauma care [4,7,12]; whilst those designated P2 (mortality 5.0%, 5.0% ICU admission rate, median ISS of 9) may appropriately be treated in second-tier centres [21]. Incorporating timing of intervention to distinguish P1 from P2 patients is also clinically meaningful in the major incident setting. Furthermore, the proportion of P1 casualties (11.3%) yielded is comparable to the 10.8–17.7% critical injury rate reported in recent UK [4,10] and international major incidents [8,29]. We recommend that Lerner's criteria [21] be employed as the gold standard for future use in post-event evaluations, research, and training purposes in the UK major incident setting. This will enable standardised assessment of major incidents and triage systems, aiding resource planning and policy refinement.

The Expectant category (also known as P1 hold, P4 or T4) has previously been defined as "casualties whose condition is so severe that they cannot survive despite the best available care and whose treatment would divert medical resources from salvageable patients who may then

be compromised [12].” However, triage category definitions proposed by Lerner did not take into account resource limitations²¹: the definition employed in this study was associated with near-universal mortality (99.1%) and assigned to patients whose injuries would be non-survivable under any circumstances (e.g. 90% burns). The Expectant category has never been assigned in UK civilian practice⁴: rationing healthcare to severely injured, living patients is ethically challenging [11,12]. We recommend that this non-resource dependant definition of the Expectant category is appropriate for academic use and that in practice, Expectant status should only be assigned by a senior clinician.

Weaknesses of this study include use of singly injured TARN patients as surrogates for those injured in major incidents, in whom outcomes (e.g. mortality) may be considerably worse. Blunt trauma predominates in this registry population: therefore, triage tool performance may not be completely generalisable to penetrating or blast mechanism incidents. Some assumptions made to facilitate retrospective triage tool application (e.g. patients meeting TARN inclusion criteria are non-ambulatory) may not hold true in real-life. Study conclusions were unlikely to be biased by patients (9.2%) excluded due to missing data. TARN inclusion criteria are biased towards capturing the severely injured, however this can be viewed as a strength since these patients are at greatest risk of adverse outcome: care received is unlikely to greatly influence outcome in the large number of “walking wounded” patients who predominate in major incidents [1,6,10]. Additionally, TARN excludes pre-hospital deaths. Given that most trauma deaths occur pre-hospital [8,10,29], our study fails to capture the main indicators of early mortality by only analysing tool performance in patients who reach hospital alive; however, those who do not make it to hospital alive under normal circumstances are even less likely to do so during a major incident, where transfer will likely be delayed due to pre-hospital resource constraints. Furthermore, our study focussed on tool performance in predicting P1 status: further research is required to evaluate tool performance in predicting non-P1 categories. Notably, triage tools commonly assign P3 status to those able to walk, however a Dutch study of an aeroplane crash involving 135 casualties revealed serious underlying injuries in 17% of ambulatory patients [20]. Finally, computed retrospective application of triage tools is used as a surrogate for EMS personnel conducting real-time triage under challenging circumstances. This does not account for human error, or the variation in tool complexity which affects their ease and accuracy of application [5,11]. However, the current study design allows the physiological discriminative capability of tools to be assessed independently of human error, as well as overcoming the challenges of conducting prospective studies during major incidents. Consequently, registry-based studies such as this are likely to form the highest level of evidence to guide major incident triage practice and policy.

In practice, policy makers must consider several factors before adopting a triage tool for widespread use [3,5,8,11]. This includes ability of the tool to differentiate categories accurately across a variety of relevant injury mechanisms [5,16]; time taken and inter-rater reliability when applied [5,8,23]; the degree of over-triage that health systems can accommodate [1,10]; and interoperability between the multiple agencies involved in major incidents [3,8,10,11]. Additionally, regular major incident training exercises have been credited for improved performance during the 2017 Manchester Arena [4] and 2015 Paris terrorist attacks [29]. Whilst simple triage algorithms have been employed since the 1980s [5,11], further work is required to determine how technology (e.g. portable device applications, electronic patient management systems, wearable devices) may be effectively incorporated into clinical use to augment triage.

In conclusion, based on current available evidence and the findings of this study, the NARU Triage Sieve used currently in UK civilian practice is not the optimal tool for major incident triage. We recommend its replacement by the BCD Triage Sieve, which may afford a

24–26% improvement in detecting patients in need of time-critical, life-saving intervention at the scene of a major incident, thus facilitating their immediate transfer to the highest tier of trauma care and maximising chances of survival. ISS > 15 correlates poorly with the need for life-saving intervention, making this a suboptimal endpoint in evaluating major incidents. We have validated a system of retrospectively assigning triage categories based on commonly described EMS and hospital interventions, resulting in clinically distinct groups of patients: we recommend its use in future evaluations of UK major incidents.

5. Author contributions

NM conducted a literature review prior to the study. SC and YX contributed equally to this study. NM, JV, AB, DB, DK, MF and GVG designed the study. NM, SC and YX accessed the database, verified the underlying data and conducted analysis. All authors contributed to data interpretation. NM wrote the initial draft of the manuscript. All authors contributed to critical revisions of subsequent manuscript drafts and approve of the final version.

Declaration of Competing Interest

The authors confirm that they have no conflicts of interest to declare.

Data sharing statement

De-identified patient data utilised for this study are proprietary to the Trauma Audit and Research Network, University of Manchester and may be requested directly from TARN[22].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.eclinm.2021.100888](https://doi.org/10.1016/j.eclinm.2021.100888).

References

- [1] Centre for Research on the Epidemiology of Disasters and UN Office for Disaster Risk Reduction. Human cost of disasters: an overview of the last 20 years 2000–2019. Emergency Events Database (EM-DAT) Centre for Research on the Epidemiology of Disasters; 2020. p. 1–17.
- [2] Moran CG, Webb C, Brohi K, Smith M, Willett K. Lessons in planning from mass casualty events in UK. *BMJ* 2017;359:j4765.
- [3] World Health Organisation. Mass casualty management systems: strategies and guidelines for building health sector capacity. Prevention HAICIV; 2007 editor.
- [4] Deeming H. The Kerslake Report: an independent review into the preparedness for, and emergency response to, the Manchester Arena attack on 22nd May 2017, 2018.
- [5] Jenkins JL, McCarthy ML, Sauer LM, et al. Mass-casualty triage: time for an evidence-based approach. *Prehospital Disaster Med* 2008;23(1):3–8.
- [6] Kahn CA, Schultz CH, Miller KT, Anderson CL. Does START triage work? An outcomes assessment after a disaster. *Ann Emerg Med* 2009;54(3):424–30 30.e1.
- [7] American College of Surgeons Committee on Trauma. Resources for Optimal Care of the Injured Patient, 2014.
- [8] Turner CD, Lockey DJ, Rehn M. Pre-hospital management of mass casualty civilian shootings: a systematic literature review. *Crit Care* 2016;20(1):362.
- [9] Frykberg ER, Tepas 3rd JJ. Terrorist bombings. Lessons learned from Belfast to Beirut. *Ann Surg* 1988;208(5):569–76.
- [10] Aylwin CJ, König TC, Brennan NW, et al. Reduction in critical mortality in urban mass casualty incidents: analysis of triage, surge, and resource use after the London bombings on July 7, 2005. *Lancet North Am Ed* 2006;368(9554):2219–25.
- [11] Lerner EB, Schwartz RB, Coule PL, et al. Mass casualty triage: an evaluation of the data and development of a proposed national guideline. *Disaster Med Public Health Preparedness* 2008;2(1):S25–34 **Suppl**.
- [12] National Ambulance Resilience Unit (NARU). Clinical guidance: medical support minimal requirements for a mass casualty incident. www.naru.org.uk; 2014. Accessed June 2020
- [13] Hodgetts TJ, Hanlan CG, Newey CG. Battlefield first aid: a simple, systematic approach for every soldier. *J R Army Med Corps* 1999;145(2):55–9.
- [14] Vassallo J, Smith J, Wallis L. Major incident triage and the implementation of a new triage tool, the MPTT-24. *J R Army Med Corps* 2018;164:103–6.

- [15] Garner A, Lee A, Harrison K, Schultz CH. Comparative analysis of multiple-casualty incident triage algorithms. *Ann Emerg Med* 2001;38(5):541–8.
- [16] Cross KP, Cicero MX. Head-to-head comparison of disaster triage methods in pediatric, adult, and geriatric patients. *Ann Emerg Med* 2013;61(6):668–76 e7.
- [17] Vassallo J, Smith JE, Wallis LA. Major incident triage and the implementation of a new triage tool. *the MPTT-24* 2018;164(2):103–6.
- [18] Bennett A. Methodologies utilized and lessons learned in high threat environments and mass casualty environments. *J High Threat Austere Med* 2019:1–7.
- [19] Office of National Statistics. Population estimates 2019. 2019. Accessed June 2020
- [20] Postma IL, Weel H, Heetveld MJ, et al. Mass casualty triage after an airplane crash near Amsterdam. *Injury* 2013;44(8):1061–7.
- [21] Lerner EB, McKee CH, Cady CE, et al. A consensus-based gold standard for the evaluation of mass casualty triage systems. *Prehospital Emergency Care* 2015;19(2):267–71.
- [22] The University of Manchester. The trauma audit and research network. Continuous. www.tarn.ac.uk. Accessed June 2020
- [23] Wolf P, Bigalke M, Graf BM, Birkholz T, Dittmar MS. Evaluation of a novel algorithm for primary mass casualty triage by paramedics in a physician manned EMS system: a dummy based trial. *Scand J Trauma Resusc Emerg Med* 2014;22:50.
- [24] Cobas MA, De la Peña MA, Manning R, Candiotti K, Varon AJ. Prehospital intubations and mortality: a level 1 trauma center perspective. *Anesth Analg* 2009;109(2):489–93.
- [25] Deakin CD, Low JL. Accuracy of the advanced trauma life support guidelines for predicting systolic blood pressure using carotid, femoral, and radial pulses: observational study. *BMJ (Clinical research ed)* 2000;321(7262):673–4.
- [26] Kelly CA, Upex A, Bateman DN. Comparison of consciousness level assessment in the poisoned patient using the alert/verbal/painful/unresponsive scale and the Glasgow Coma Scale. *Ann Emerg Med* 2004;44(2):108–13.
- [27] Hollander M, Wolfe DA, Chicken E. Nonparametric statistical methods. Wiley; 2013.
- [28] Sun X, Xu W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Process Lett* 2014;21(11):1389–93.
- [29] Hirsch M, Carli P, Nizard R, et al. The medical response to multisite terrorist attacks in Paris. *Lancet* 2015;386:2535–8 10012.
- [30] Moran CG, Lecky F, Bouamra O, et al. Changing the system - major trauma patients and their outcomes in the NHS (England) 2008–17. *Eclin Med* 2018;2:13–21.
- [31] Baxt WG, Upenieks V. The lack of full correlation between the injury severity score and the resource needs of injured patients. *Ann Emerg Med* 1990;19(12):1396–400.