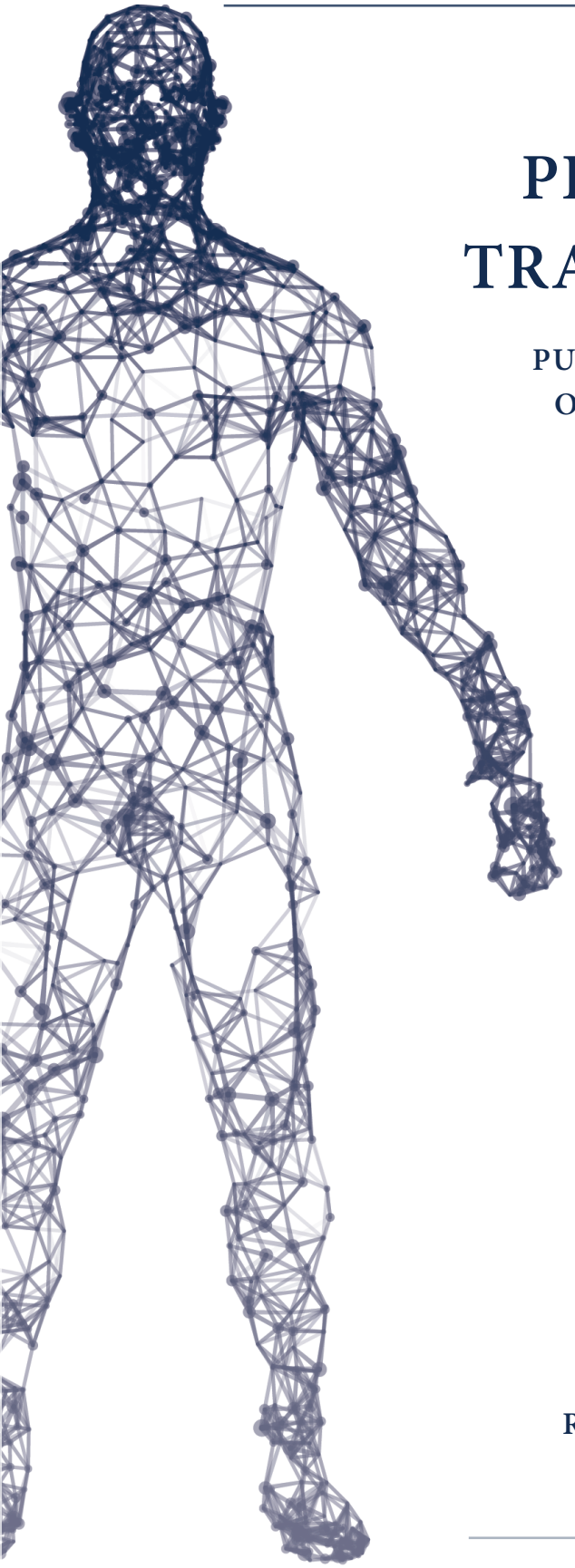


PRE-HOSPITAL TRAUMA SYSTEMS

PUSHING THE BOUNDARIES OF PRECISION MEDICINE

ROGIER VAN DER SLUIJS

PRE-HOSPITAL TRAUMA SYSTEMS

PUSHING THE BOUNDARIES OF PRECISION MEDICINE

ROGIER VAN DER SLUIJS

Pre-hospital Trauma Systems,
Pushing the Boundaries of Precision Medicine

Thesis, Utrecht University, The Netherlands

ISBN/EAN: 9789464023565

Design: wenz iD, Wendy Schoneveld
Printed by: Gildeprint Drukkerijen B.V., Enschede

The research described in this thesis was financially supported by:
A grant from ZonMw, the Dutch Organization for Health Research (80-84300-98-18555) and Development and NVIDIA Corporation.

Printing of this thesis was financially supported by:
Netwerk Acute Zorg Limburg

Pre-hospital Trauma Systems

Pushing the Boundaries of Precision Medicine

Pre-hospitale trauma systemen:
Nieuwe toepassingen van precisiegeneeskunde
(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
woensdag 8 juli des ochtends te 11.00 uur

door

ROGIER VAN DER SLUIJS

geboren op 27 juli 1991
te Utrecht

PROMOTOREN

Prof. dr. L.P.H. Leenen

Prof. dr. M. Poeze

COPROMOTOR

Dr. M. van Heijl

Voor mijn ouders

CONTENTS

CHAPTER I

General introduction

The current goal of field triage is embodied by the maxim "*getting the right patient, to the right place, in the right time*".[1] The principles of this tripartite strategy can be traced back to the revolutionary ideas of two military surgeons that served during the French Revolution and in the Napoléonic Wars: Pierre-François Percy (1754 – 1825) and Dominique-Jean Larrey (1766 – 1842).[2] Their beliefs, innovations, and methods defined field triage and form the backbone of modern trauma systems.

Evolution of Field Triage

Until well after the Middle Ages, wounded soldiers were often left on the battlefield for days.[3] Potentially salvageable patients frequently died of wounds, cold, or simply thirst.[3,4] Injured patients were transported to field hospitals only after a long delay or whenever a ceasefire was established. In 1792, Larrey joined the Army of the Rhine as an assistant surgeon, and soon recognized the need for swift transportation of wounded soldiers from active battlefield.[4] Larrey, who later became Surgeon in Chief of the French army, completely re-organized the field medical services and designed *ambulance volantes* to rapidly transport wounded soldiers from active battlefield to nearby field hospitals.[2-4] These *ambulance volantes* (i.e., horse-drawn wagons) were first tested in the Battle of Metz (1793) and proved to be pivotal to reduce the time between injury and definitive surgical treatment.[4,5]

The first concept of *triàge* was supposedly developed a few years later during Napoléon Bonaparte's expeditions in Egypt and Syria (1797 – 1801).[4] Many French soldiers died during these difficult campaigns and it was allegedly Percy who developed a system to prioritize treatment of the sick and injured soldiers that could potentially return to the battlefield.[4] This early concept of Napoléonic triage was later refined by Larrey, who pursued a more philanthropic approach.[4] Eventually, Larrey established a new triage paradigm that categorized patients based on three grades of injury severity: dangerously wounded, moderately wounded, and mildly wounded.[4]

Evolution of Trauma Systems

It took several decades before these revolutionizing concepts of ambulance systems and triage were re-applied, most notably in the American Civil War (1861 – 1865).[2,4] Presumably the first trauma manual that documented a triage strategy, aid stations, rapid transport to hospitals, and care of injured patients, was conceived during this war.[6] The many lessons learned during the American Civil War, World War I, and World War II set the stage for modern trauma systems. In 1966, a landmark paper entitled *Accidental Death and Disability: The Neglected Disease of Modern Society* was published by the American National Academy of Sciences and the National Research Council.[7] The outgrowth of this report was the development of regional Emergency Medical Services

(EMS), emergency telephone systems, paramedic training programs, and systems of trauma care.[8]

The American College of Surgeons Committee on Trauma (ACSCOT) recognized the need for a nationwide standard for evaluation of trauma care and published a guidebook entitled *Optimal Hospital Resources for Care of the Seriously Injured* in 1976.[9] This was the first manual that comprised criteria for trauma center designation.[8] Later revisions of this document refined these criteria, emphasized the importance of ongoing quality monitoring, and focused on the principles of developing all-encompassing trauma systems.[10] These so-called *inclusive trauma systems* (i.e., regionalized systems of trauma care) incorporate all elements needed for optimal care along the continuum of trauma care.

In 1987, a thesis on trauma care evaluation in the Netherlands appeared that advocated for regionalized trauma care, designation of trauma centers, and the use of field triage protocols, in line with the recommendations of the ACSCOT.[11,12] The Dutch Medical Health Inspectorate subsequently investigated the continuum of emergency care and underlined the need for reorganization in a 1994 report.[13] The Dutch Society of Trauma Surgery responded in 1997 with a manifest that formed the blueprint for a regionalized system of trauma care that was finally instituted in 1999.[14]

At present, the Netherlands is divided into 11 geographically defined inclusive trauma regions. Each region has a single coordinating level-I trauma center and encompasses multiple level-II and level-III trauma centers. The designated levels reflect the role of trauma-receiving hospitals within the trauma system.[10] Level-I trauma centers are considered to be higher-level trauma centers (similar to level-I and level-II in the USA) with sufficient expertise and resources to treat the most severely injured patients. Lower-level trauma centers (i.e., level-II and level-III), on the other hand, were established to provide optimal care for moderately and mildly injured patients in a cost-effective manner. This systematic and regionalized approach to trauma care was evaluated in multiple studies, including a study evaluating the Dutch trauma system, and was reported to decrease mortality rates and injury-related re-admissions.[15-18]

Modern Field Triage

The original concept of triage relates to mass casualty situations, typically during active warfare.[5] The onset of civilian systems of trauma care introduced three new phases of triage: determination of ambulance priority by the EMS dispatch center, field triage at the scene of injury by EMS professionals, and triage on arrival at the emergency department. This thesis focuses on field triage of trauma patients in inclusive trauma systems, although abstract concepts may very well apply to other phases of triage, to trauma team activation, and to other clinical domains.

Field triage is a diagnostic or prognostic strategy that consists of multiple steps: (1) assessment of injury severity, (2) estimation of a patient's resource-need, (3) the determination of hospitals that match the patient's needs, and (4) the assessment of available hospitals, given trauma center proximity, trauma center capacity, and patient acuity. The final decision on the initial transport destination needs to be determined in light of these steps by EMS professionals at the scene of injury.

Triage Tools

One key element of the triage strategy is a pre-hospital triage protocol that is designed to aid EMS professionals in their assessment of injury severity, the estimation of patients' suspected resource-need, and the choice of the initial transportation destination. Various field triage tools have been developed in the past to assist EMS professionals at the scene of injury.[19] The Trauma Score, Revised Trauma Score, Pre-hospital Index, and the Circulation, Respiration, Abdomen, Motor, and Speech scale were all innovative tools developed in the 1980s but lacked the predictive ability to classify injury severity accurately.[20-24] A novel idea, ahead of its time, was proposed by Baxt in 1991.[25] Baxt uncovered flaws in the evaluation of the accuracy of triage tools and proposed a new resource-based triage rule.[25,26] Moreover, Baxt recognized the need for more complex prediction models, as appears from his pioneering work on artificial neural networks to detect acute myocardial infarctions in the emergency department.[27,28] Most of these tools, including the resource-based triage rule, fell into disuse. Many inclusive trauma systems in the USA implemented a version of the five-times revised Field Triage Decision Scheme that was originally developed by the ACSCOT in 1986.[10] Nowadays, few different flavors of triage protocols exist, and most wholly or in part rely on triage criteria originating from the Field Triage Decision Scheme, such as the National Protocol of Ambulance Services (Choice of Hospital; in Dutch, *Landelijke Protocol Ambulancezorg – Keuze ziekenhuis*) of the Dutch Institute of Ambulance Care.[29]

Evaluation of Regionalized Trauma Care

Mistriage can be detrimental in mature trauma systems with a high degree of resource centralization. Undertriage – transporting patients requiring specialized trauma care to lower-level trauma centers – is associated with increased mortality rates and could potentially lead to avoidable life-long disabilities.[30] In contrast, overtriage – transporting patients without the need for specialized trauma care to higher-level trauma centers – is associated with excessive costs and overutilization of scarce resources.[31] Generally, priority has been given to reducing undertriage rates. The ACSCOT recommends that trauma systems should attain an undertriage rate of less than 5%.[10] The Dutch Healthcare Institute published similar guidelines in 2015 that required the inclusive trauma system to reduce undertriage to less than 10%.[32] A systematic review in 2018 unveiled that not a single

inclusive trauma system worldwide was able to attain the recommended undertriage rates in adults, while preserving generally accepted overtriage rates of up to 35%.[10,33] In addition, another systematic review was unable to identify triage tools with acceptable predictive ability during external validation.[19]

Thesis Outline

Present-day field triage strategies are dichotomous: patients are either considered to be severely injured or not, patients are in need of specialized trauma care or not, and trauma centers either have sufficient resources to treat severely injured patients (i.e., higher-level) or not (i.e., lower-level). The aim of this thesis is to replace this overly simplistic view of reality with a personalized strategy to evaluate and optimize field triage in inclusive trauma systems. The research questions covered by each chapter are outlined in the Table. Four main topics regarding the evaluation and optimization of field triage strategies are addressed in this thesis:

I. Evaluation of pediatric and adult pre-hospital trauma triage
II. Construction of a population-based cohort to continuously monitor, evaluate, and investigate triage accuracy
III. Development and validation of prediction models to identify patients in need of specialized trauma care
IV. Future perspectives of field triage in inclusive trauma systems

Table The 15 Study Questions Addressed in this Thesis

Chapter	Study question
2	What is the accuracy of pre-hospital trauma triage based on the initial transportation destination in a pediatric population around the world?
2	Which field triage tools that aid Emergency Medical Services professionals in the determination of the initial transportation destination were externally validated in a pediatric population and what is their diagnostic accuracy?
3	What is the accuracy of pre-hospital trauma triage based on the initial transportation destination in a pediatric population in the Netherlands?
3	How do contemporary field triage tools perform in terms of accuracy in a pediatric population based on an anatomical and a resource-based reference standard?
4	What is the accuracy of pre-hospital trauma triage based on the initial transportation destination in an adult population in the Netherlands?
4	What is the accuracy of the National Protocol of Ambulance Services to select adults in need of specialized trauma in the pre-hospital setting based on an anatomical reference standard?
5	How can we establish a prospective, open, and voluminous cohort to allow continuous monitoring of the pre-hospital trauma system?
5	What is the performance of a prediction model developed to select trauma patients from unfiltered pre-hospital electronic health records?
5	What is the performance of a prediction model developed to link pre-hospital electronic health records to in-hospital patient outcomes collected by the Dutch National Trauma Registry?
6	What is predictive performance and external validity of a model to select severely injured patients in the pre-hospital setting based on an anatomical reference standard?
7	What would be an adequate strategy to develop and validate a prediction model to identify patients in need of specialized trauma care?
8	What is the predictive ability and external validity of novel prediction models to predict injury severity and patients' resource use?
8	What is the net benefit of the new prediction models compared to contemporary field triage tools?
9	How could the impact of the *Trauma Triage App* be assessed in daily practice?
10	How can we optimize and evaluate field triage in inclusive trauma systems in the future?

PART I: EVALUATION OF FIELD TRIAGE

Recently, a series of three systematic reviews was published that summarized evidence on destination-based field triage accuracy, protocol accuracy, and the compliance of EMS professionals with field triage protocols.[19,33,34] These reviews concluded that (i) there was a lack of methodologically sound studies that investigated destination-based triage accuracy, that (ii) there was no single trauma system worldwide with both acceptable undertriage and overtriage rates, (iii) nor a triage tool that was able to attain such rates, and that (iv) conformity to these tools was generally low. Inclusion in these reviews was restricted to adults and the applicability of these findings remains unclear in pediatric populations. This gap of knowledge was a matter of serious concern given that injuries continue to be a leading cause of death and disability among children.[35] In *Chapter II* we summarize evidence on triage accuracy and protocol accuracy in children by means of a systematic review of the literature. In *Chapter III* we subsequently evaluate various aspects of the pediatric field triage strategy in a Dutch multi-site observational cohort study.

The bulk of literature on systems of regionalized trauma care describes the American setting since few European studies on this subject exist.[19] *Chapter IV* investigated the triage strategy in a single inclusive trauma region in the Netherlands. Moreover, similar to *Chapter III*, it externally validates the Dutch National Protocol of Ambulance Services that is actively used to guide EMS professionals on the allocation of patients in the Netherlands.

PART II: THE TRAUMA CONTINUUM OF CARE COHORT

Evaluation of pre-hospital systems requires the availability of pre-hospital data and data on patient outcomes. The Dutch National Trauma Registry (in Dutch, *Landelijke Traumaregistratie*) is an excellent source of patient outcomes that includes prospectively collected data from all hospitals with trauma-receiving emergency departments in the Netherlands. Pre-hospital data, however, are not fully covered in this registry and a consecutive series of patients with pre-hospital data was lacking. *Chapter V* describes the methodology used to construct a multi-site observational cohort of trauma patients that includes both pre-hospital and in-hospital data. In this chapter, we describe the Trauma Triage Continuum of Care Cohort that enables researchers to continuously study the effects of pre-hospital trauma care and decision-making on in-hospital patient outcomes.

PART III: PERSONALIZED FIELD TRIAGE

Contemporary triage tools, such as the Dutch National Protocol of Ambulance Services, are mostly based on similar criteria to the Field Triage Decision Scheme of the ACSCOT. Based on these flowcharts, transport to a higher-level trauma center is advised if a patient fulfills one or more physiologic or anatomic criteria. Predictive ability of such flowcharts is generally suboptimal and particularly poor in the heterogeneous trauma population.[19] *Chapter VI* describes the development and validation of a prediction model to identify severely injured patients at the scene of injury. In *Chapter VII* we outline a robust model-development-and-validation strategy to identify patients in need of specialized trauma care during field triage. We implement this methodology in *Chapter VIII* to construct two models that aim to predict well-calibrated probabilities for early critical-resource use and presence of severe injuries. In this chapter we intended to derive prediction models that require minimal human input, work well under the circumstances of incomplete data, and dynamically update predictions whenever new data become available. In *Chapter IX* we propose the design of a stepped-wedge cluster randomized trial to evaluate the impact of a new smartphone-based triage tool on mistriage rates.

PART IV: DYNAMIC TRAUMA SYSTEMS

In the final part of this thesis, we present our opinion on the future perspectives of field triage in inclusive trauma systems. *Chapter X* describes imperfections in the current methodology used to evaluate pre-hospital trauma systems and proposes a new personalized approach to optimize field triage in the future.

REFERENCES

1. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Dec 15, 2019).

2. Baker D, Cazalaa JB, Carli P. Resuscitation great. **Larrey and Percy--a tale of two barons.** Resuscitation 2005; 66(3): 259-62.

3. Pollock A. Historical perspectives in the ambulance service. In: Wankhade P, Mackaway-Jones K, eds. **Ambulance services: leadership and management perspectives.** Switzerland: Springer International; 2015.

4. Nakao H, Ukai I, Kotani J. **A review of the history of the origin of triage from a disaster medicine perspective.** Acute Med Surg 2017; 4(4): 379-84.

5. Robertson-Steel I. **Evolution of triage systems.** Emerg Med J 2006; 23(2): 154-5.

6. American College of Surgeons. **A Brief History of Trauma Systems.** https://www.facs.org/quality-programs/trauma/tqp/systems-programs/trauma-series/part-i (accessed Dec 15, 2019).

7. National Academy of Sciences. **Accidental death and disability: the neglected disease of modern society.** Washington DC, 1966.

8. Coimbra R, Hoyt DB, Bansal V. **Trauma Systems, Triage and Transport.** In: Mattox LK, Moore EE, Feliciano DV, eds. Trauma. 7 ed. New York: McGaw-Hill; 2013.

9. American College of Surgeons Committee on Trauma. **Optimal hospital resources for care of the seriously injured.** Bull Am Coll Surg 1976; 61(9): 15-22.

10. American College of Surgeons Committee on Trauma. **Resources for the optimal care of the injured patient.** Chigago, IL, 2014.

11. Draaisma JM, de Haan AF, Goris RJ. **Preventable trauma deaths in The Netherlands--a prospective multicenter study.** J Trauma 1989; 29(11): 1552-7.

12. Draaisma JMT. **Evaluation of Trauma Care.** Nijmegen: Radboud University; 1987.

13. Geneeskundige Inspectie van de Volksgezondheid. **De keten rammelt: onderzoek naar de afstemming binnen de spoedeisende hulpverlening**; 1994.

14. Traumatologie NVvH. **Trauma zorg; onze zorg! Blauwdruk voor een verbeterde traumazorg in Nederland;** 1997.

15. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

16. Celso B, Tepas J, Langland-Orban B, et al. **A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems.** J Trauma 2006; 60(2): 371-8.

17. Curtis K, Chong S, Mitchell R, Newcombe M, Black D, Langcake M. **Outcomes of severely injured adult trauma patients in an Australian health service: does trauma center level make a difference?** World J Surg 2011; 35(10): 2332-40.

18. Twijnstra MJ, Moons KG, Simmermacher RK, Leenen LP. **Regional trauma system reduces mortality and changes admission rates: a before and after study.** Ann Surg 2010; 251(2): 339-43.

19. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld RA, Leenen LPH, van Heijl M. **Accuracy of prehospital triage protocols in selecting severely injured patients: A systematic review.** J Trauma Acute Care Surg 2017; 83(2): 328-39.

20. Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. **Trauma score.** Crit Care Med 1981; 9(9): 672-6.

21. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. **A revision of the Trauma Score.** J Trauma 1989; 29(5): 623-9.

22. Gormican SP. **CRAMS scale: field triage of trauma victims.** Ann Emerg Med 1982; 11(3): 132-5.

23. Koehler JJ, Baer LJ, Malafa SA, Meindertsma MS, Navitskas NR, Huizenga JE. **Prehospital Index: a scoring system for field triage of trauma victims.** Ann Emerg Med 1986; 15(2): 178-82.

24. Baxt WG, Berry CC, Epperson MD, Scalzitti V. **The failure of prehospital trauma prediction rules to classify trauma patients accurately.** Ann Emerg Med 1989; 18(1): 1-8.

25. Baxt WG, Jones G, Fortlage D. **The trauma triage rule: a new, resource-based approach to the prehospital identification of major trauma victims.** Ann Emerg Med 1990; 19(12): 1401-6.

26. 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.

27. Baxt WG. **Use of an artificial neural network for the diagnosis of myocardial infarction.** Ann Intern Med 1991; 115(11): 843-8.

28. Baxt WG, Skora J. **Prospective validation of artificial neural network trained to identify acute myocardial infarction.** Lancet 1996; 347(8993): 12-5.

29. Ambulancezorg Nederland**. Landelijk Protocol Ambulancezorg (versie 8.1).** Zwolle, 2016. https://www.ambulancezorg.nl/themas/kwaliteit-van- zorg/protocollen-en-richtlijnen/landelijk-protocol-ambulancezorg (accessed Dec 15, 2019).

30. Staudenmayer K, Weiser TG, Maggio PM, Spain DA, Hsia RY. **Trauma center care is associated with reduced readmissions after injury.** J Trauma Acute Care Surg 2016; 80(3): 412-6; discussion 6-8.

31. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

32. Zorginstituut Nederland. **Spoed moet goed: indicatoren en normen voor zes spoedzorgindicaties**; 2015.

33. van Rein EAJ, van der Sluijs R, Houwert RM, et al. **Effectiveness of prehospital trauma triage systems in selecting severely injured patients: Is comparative analysis possible?** Am J Emerg Med 2018; 36(6): 1060-9.

34. van Rein EAJ, van der Sluijs R, Raaijmaakers AMR, Leenen LPH, van Heijl M. **Compliance to prehospital trauma triage protocols worldwide: A systematic review.** Injury 2018; 49(8): 1373-80.

35. WHO. **World report on child injury prevention.** Geneva, 2008. https://www.who.int/violence_injury_prevention/child/injury/world_report/en/ (accessed Dec 15, 2019).

PART I

Evaluation of field triage

CHAPTER II

Accuracy of pediatric trauma field triage: a systematic review

JAMA Surgery

Rogier van der Sluijs; Eveline A.J. van Rein;
Joep G.J. Wijnand; Luke P.H. Leenen; Mark van Heijl

ABSTRACT

Importance

Field triage of pediatric patients with trauma is critical for transporting the right patient to the right hospital. Mortality and lifelong disabilities are potentially attributable to erroneously transporting a patient in need of specialized care to a lower-level trauma center.

Objective

To quantify the accuracy of field triage and associated diagnostic protocols used to identify children in need of specialized trauma care.

Evidence Review

MEDLINE, Embase, PsycINFO, and Cochrane Register of Controlled Trials were searched from database inception to November 6, 2017, for studies describing the accuracy of diagnostic tests to identify children in need of specialized trauma care in a pre-hospital setting. Identified articles with a study population including patients not transported by Emergency Medical Services were excluded. Quality assessment was performed using a modified version of the Quality Assessment of Diagnostic Accuracy Studies–2.

Findings

After deduplication, 1430 relevant articles were assessed, a full-text review of 38 articles was conducted, and five of those articles were included. All studies were observational, published between 1996 and 2017, and conducted in the United States, and data collection was prospective in one study. Three different protocols were studied that analyzed a combined total of 1222 children in need of specialized trauma care. One protocol was specifically developed for a pediatric out-of-hospital cohort. The percentage of pediatric patients requiring specialized trauma care in each study varied between 2.6% (110 of 4197) and 54.7% (58 of 106). The sensitivity of the pre-hospital triage tools ranged from 49.1% to 87.3%, and the specificity ranged from 41.7% to 84.8%. No pre-hospital triage protocol alone complied with the international standard of 95% or greater sensitivity. Undertriage and overtriage rates, representative of the quality of the full diagnostic strategy to transport a patient to the right hospital, were not reported for inclusive trauma systems or Emergency Medical Services regions.

Conclusions and Relevance

It is crucial to transport the right patient to the right hospital. Yet the quality of the full diagnostic strategy to determine the optimal receiving hospital is unknown. None of the investigated field triage protocols complied with current sensitivity targets. Improved

efforts are needed to develop accurate child-specific tools to prevent undertriage and its potential life-threatening consequences.

INTRODUCTION

Injury is a leading cause of death and disability among children worldwide.[1] Field triage in inclusive trauma systems is critical to get the right patient to the right hospital to achieve optimal patient outcomes. Transporting an injured child in need of specialized trauma care to a lower-level, non-pediatric trauma center is considered undertriage and is associated with higher mortality rates.[2-5] Conversely, transporting patients without need for specialized trauma care to higher-level trauma centers (overtriage) results in overuse of valuable trauma resources and increased costs.[6]

It is crucial that Emergency Medical Services (EMS) professionals at the scene correctly determine the definitive care facility to prevent delay of care and to avoid inter-hospital transfers.[7,8] A multitude of comparable triage protocols that predict need of specialized trauma care have been developed to aid decision-making during field triage. However, most of these protocols were developed for adults or for different settings, leaving it unclear whether test performance is upheld for injured children triaged by EMS professionals.[9,10]

Field triage research focuses on the full diagnostic strategy used to determine which facility the patients should be directly transported to, with undertriage and overtriage as key quality metrics. Triage protocols to predict injury severity are crucial elements of this strategy, and the performance quality of these protocols can be expressed in terms of sensitivity and specificity. The objective of this systematic review is to summarize evidence on triage accuracy during field triage among children suspected of injury.

METHODS

Data Sources and Search

The review was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines.[11] Studies were searched in MEDLINE, PsycINFO, Embase, and the Cochrane Libraries from database inception until November 6, 2017. Search terms included pediatric trauma (study population), triage protocols (index tests), accuracy (outcome), and field triage (setting; see Appendix). The reference lists of studies reviewed during full-text article assessment were checked for additional eligible studies.

Eligibility Criteria

Eligible were cross-sectional studies, cohort studies, and randomized clinical trials of patients suspected of injury who were evaluated in a pre-hospital setting and transported by EMS. We included studies using composite outcome measures of early critical-resource use or surrogate markers for severe injury as reference standards for test performance. Studies including all patients presenting to the emergency department regardless of transportation type (i.e., including private transportation) were excluded. Studies were excluded when accuracy metrics for children (0 to <18 years of age) were not reported separately and could not be calculated. No language, publication date, or publication status restrictions were imposed.

Outcomes

Performance measures were calculated for each index test with its corresponding reference standard. Sensitivity and specificity were calculated along with 95% confidence intervals for triage protocols. Undertriage (proportion of severely injured patients initially transported to a lower-level trauma center) and overtriage (proportion of patients without severe injuries initially transported to a higher-level trauma center or PTC) were calculated to define the overall triage accuracy for a region. Confidence intervals incorporating between and within-variance could not be calculated when studies used multiple imputation to address missing values. Pooled estimates were considered but were assumed trivial owing to heterogeneous index tests and the limited number of studies covering this subject.

Study Selection and Data Extraction

The eligibility assessments were performed by two reviewers (RvdS and JGJW) in a standardized manner. Non-duplicate records were screened by title and abstract, after which the full text of the remaining records was reviewed. The same two extracted relevant study characteristics using a data extraction template based on the Standards for Reporting Diagnostic Accuracy checklist.[12] Information was extracted on: (1) design, setting, and inclusion criteria; (2) index tests and reference standards; (3) accuracy metrics calculated from tables (2×2) derived for each index test and its corresponding reference standard. Assessment of risk of bias was performed by these same two reviewers using a modified version of the Quality Assessment of Diagnostic Accuracy Studies-2 tool.[13] Any discrepancies in study selection, data extraction, and quality ratings were resolved by consensus.

RESULTS

Search

The initial search yielded 1429 unique records (Figure). After screening on title and abstract, 37 records remained for full-text review. One additional record was identified through a survey of the reference lists in the reviewed studies. Five studies met our pre-specified eligibility criteria.

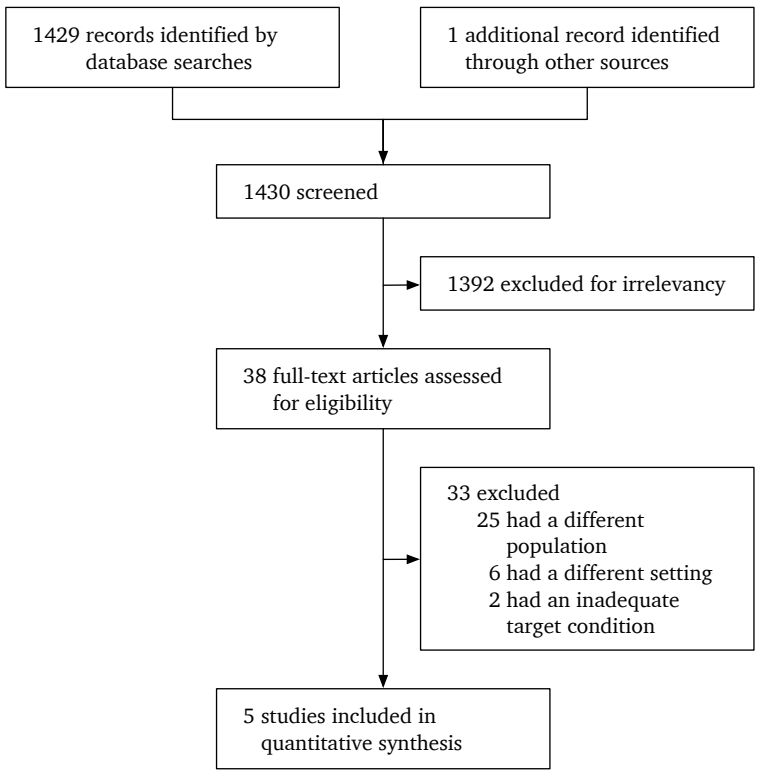


Figure Study Selection

Study Characteristics

All included studies investigated the accuracy of a single field triage protocol. None of the studies reported regional triage accuracy based on the initial destination facility. Most triage protocols were applied retrospectively to pre-hospital factors collected by run-reports, existing hospital databases, and registries. The included studies were published between 1996 and 2017. All studies were conducted in the United States.

Table 1 Summary of Included Studies on Field Triage Protocols in a Pediatric Population

Source	Study design and setting	Age, y	Index test	Reference standard
Johnson et al,[18] 1996	Observational study with retrospective data collection; multiple local trauma centers in nine Florida counties in the United States in 1993	≤15	Pediatric Trauma Triage Checklist	MacKenzie algorithm
Phillips et al,[17] 1996	Observational study with retrospective data collection; acute care facilities in nine Florida counties in the United States in 1991	≤14	Trauma Scorecard	MacKenzie algorithm
Newgard et al,[16] 2011	Observational study with retrospective data collection; seven sites with 122 acute care hospitals, facilitated by 94 EMSs across the western United States evaluated from 2006 to 2008	≤17	Multiple adaptations of the Field Triage Decision Scheme (2006)	Injury Severity Score ≥16
Lerner et al,[14] 2016	Prospectively collected data in an observational study; three pediatric trauma centers across the United States were involved from 2009 to 2012	≤15	Physiologic criteria of the Field Triage Decision Scheme (2011)	Intensive care unit admission, death, or non-orthopedic surgery ≤24 h
Newgard et al,[15] 2016	Observational study with retrospective data collection; acute care facilities facilitated by 44 EMSs in seven counties in Oregon and Washington in the United States in 2011	≤14	Multiple adaptations of the Field Triage Decision Scheme (2006)	Injury Severity Score ≥16

Abbreviations: EMS, *Emergency Medical Service*.

One study collected data in a prospective manner by interviewing EMS professionals in the emergency department.[14] Four studies investigated (a part of) the protocol used in daily practice in the study region.[14-17] A new protocol was virtually tested on available data in one study.[18] The percentage of severely injured patients in each study was between 2.6% (110 of 4197 patients) and 54.7% (58 of 106 patients). The sensitivity of pre-hospital triage tools in each study ranged from 49.1% to 87.3%, and the specificity ranged from 41.7% to 84.8%.

Quality Assessment

An assessment of the risk of bias is presented in Table 2. Patient selection quality ranged from satisfactory to poor. Patients were sometimes unnecessarily excluded owing to missing values[17], an inability to match pre-hospital data to hospital or registry records[16,17], or the unavailability of research coordinators[14], leading to non-consecutive and non-random samples. Conduct of the index test was often poorly described and not according to daily practice. In two studies, clinical factors were used as surrogates for missing pre-hospital factors, giving rise to biased test accuracy.[14,17] It is unclear whether the MacKenzie

Table 2 Critical Appraisal of the Included Articles

	Risk of bias				Applicability concerns		
Source	**Patients**	**Index test**	**Reference standard**	**Flow and timing**	**Patients**	**Index test**	**Reference standard**
Johnson et al,[18] 1996	+	+	−	+	−	+	?
Phillips et al,[17] 1996	−	−	−	+	−	+	?
Newgard et al,[16] 2011	−	+	−	+	+	+	+
Lerner et al,[14] 2016	−	−	+	−	−	+	?
Newgard et al,[15] 2016	+	+	−	+	+	+	+

Symbols: +, low suspicion of bias; −, potential bias; ?, insufficient information.

algorithm, and early critical-resource use are applicable to identifying patients in need of specialized trauma care.[19] Mapping functions were used in three studies to generate Injury Severity Scores (ISSs) or Abbreviated Injury Scale scores from International Classification of Diseases (ninth revision) codes, and missing reference standards were imputed, leading to an imperfect conduct of the reference standard and potential misclassification.[16-18,20] The inclusion of patients transported only to trauma centers raised applicability concerns regarding patient selection in two studies.[14,18] It is unlikely that test accuracy of patients transported only using advanced life support extrapolates to the complete pediatric out-of-hospital population as defined in the review question.[18]

Trauma Triage Protocols

The accuracy metrics of pre-hospital triage protocols are given in Table 3. The Pediatric Trauma Triage Checklist (PTTC) is an adaptation of the Pediatric Trauma Score (PTS) that was designed to make the PTS more user-friendly.[18] The original PTS consists of the assessment of six anatomic or physiologic components, including airway, systolic blood pressure, level of consciousness, fractures, and cutaneous injuries. Each component is assigned a value of −1, 1, or 2 and a cumulative score is calculated.[21] The PTS was originally developed with in-patient data and was aimed at predicting injury severity and mortality. Early studies concluded that it was relatively hard to calculate and components had little or no meaning to EMS professionals.[18,22] The PTTC modified the component criteria to make them clearer and easier to use. In addition, checkboxes were introduced for each component to eliminate the need of calculating a score. Each item is color-coded,

and one red box or two blue boxes indicate that transport to a specialized trauma center is required. In one study, the sensitivity of the PTTC was 86.2%, with a specificity of 41.7%.[18]

The Trauma Scorecard has been used for adult field triage in Florida since 1990, however, no uniform guidelines for pediatric patients existed at that time.[17] Components of the Trauma Scorecard are systolic blood pressure, respiratory rate, Glasgow Coma Scale, burns, paralysis, ejection from vehicle, amputation proximate to wrist or ankle, and penetrating injury. The ability of this adult-specific protocol to predict pediatric injury severity was investigated, and the reported sensitivity was 66.7%, with a specificity of 84.8%.[17]

The Field Triage Decision Scheme (FTDS) was established in 1986 by the American College of Surgeons Committee on Trauma (ACSCOT).[23] Modified versions of the FTDS have appeared at regular intervals. The protocol consists of four triage steps: physiologic criteria, anatomic criteria, mechanism of injury, and special patient or system considerations. The patients should be transported to the highest level of care available

Table 3 Accuracy of Pediatric Trauma Field Triage Tools

		Outcome, No.				Performance Measure, % (95%-CI)	
Index test	**Positive RS, no. (%)**	**TP**	**FN**	**FP**	**TN**	**Sensitivity**	**Specificity**
Pediatric Trauma Triage Checklist	58	50	8	28	20	86.2%	41.7%
	(54.7%)					(74.8 – 93.1)	(28.8 – 55.7)
Trauma Scorecard	78	52	26	217	1210	66.7%	84.8%
	(5.2%)					(55.6 – 76.2)	(82.8 – 86.6)
Multiple adaptations of FTDS 2006	697	586	111	4763	9414	84.1%	66.4%
	(4.7%)					(81.1 – 86.6)	(65.6 – 67.2)
Physiologic criteria of FTDS 2011	279	137	142	935	4380	49.1%	82.4%
	(5.0%)					(43.3 – 54.9)	(81.4 – 83.4)
Multiple adaptations of FTDS 2006	110 (2.6%)	96	14	844	3243	87.3%	79.3%
						(79.6 – 92.4)	(78.1 – 80.6)

Abbreviations: CI, *Agresti-Coull confidence interval*; FN, *false negative;* FP, *false positive*; FTDS, *Field Triage Decision Scheme*; RS, *reference standard;* TN, *true negative;* TP, *true positive*.

in a trauma system when anatomic or physiologic criteria are fulfilled. Physiologic and anatomic components are the Glasgow Coma Scale, systolic blood pressure, respiratory rate, penetrating injury, flail chest, two or more proximal long-bone fractures, crushed extremity, amputation proximate to wrist or ankle, pelvic fracture, skull fracture, and paralysis. Special considerations include EMS professional judgment and patient age (children should be triaged preferentially to pediatric-capable trauma centers). The FTDS was evaluated in three studies.[14-16] One study only evaluated the physiologic criteria of the 2011 version[14]. Two studies evaluated all criteria (including mechanism of injury and special considerations) of the FTDS 2006 version (or slightly different versions).[15,16] These studies used the reference standard (ISS ≥16) as suggested by the ACSCOT.[24] The physiologic criteria had a sensitivity of 49.1% and a specificity of 82.4%. The full decision scheme had a sensitivity ranging from 84.1% to 87.3% and a specificity of 66.4% to 79.3%.

DISCUSSION

This systematic review included five studies with a combined number of 1222 pediatric patients requiring specialized trauma care who were classified using three different reference standards. Maximum sensitivity of all evaluated protocols was 87.3%. These findings are important because of the potentially life-threatening consequences of erroneous field triage.

In 1976, the ACSCOT published criteria for categorizing hospitals according to their resources and expertise to treat traumatic injuries. Regionalization of trauma care is often based on these criteria, and evidence is suggestive of decreased mortality rates compared with exclusive systems.[25,26] The ACSCOT recommends an undertriage rate of less than 5% in inclusive trauma systems.[24] The present review showed that no existing protocol attained greater than 95% sensitivity to achieve this goal in a pediatric pre-hospital population. This finding is congruent with a recent review of protocol accuracy in a slightly broader out-of-hospital population of adults.[10]

Triage protocols are a single component of the diagnostic strategy used to determine the optimal definitive care facility. Besides triage protocols, this diagnostic strategy often includes EMS professional judgment, trauma center proximity, and regional agreements, and is dependent on trauma center capacity. This strategy ultimately leads to an optimal or suboptimal choice of receiving hospital. In the present review, no study evaluated regional undertriage and overtriage rates. Evidence suggests that undertriage rates are greater than 20% for children 0 to 10 years of age and 11 to 20 years of age, but no exact numbers could be computed.[27] Pediatric undertriage rates of greater than 20% are also

reported for patients admitted to emergency departments in the United States.[28,29] However, one-third of these injured patients used private transportation, and accuracy cannot be extrapolated to EMS professional-triaged patients.[30]

The PTTC was the only child-specific triage protocol included. The PTS, on which the PTTC was based, is perhaps the most studied child-specific protocol, but no study using the PTS met our inclusion criteria. Application of adult physiologic criteria to children will presumably lead to misclassification of the need for specialized trauma care.[14,31,32] Even the use of child-specific cut-points for physiologic criteria will likely result in false predictions, owing to a great variability in reference ranges across childhood.[33] In addition to differences among physiologic variables, pediatric injury patterns and mechanism of injury markedly differ from their adult counterparts. School-aged children are at greatest risk of traumatic brain injuries, mostly owing to motor vehicle crash-related trauma, whereas toddlers and preschoolers are most commonly at risk of falls. Developing a diagnostic test with acceptable accuracy across all age ranges is consequently challenging.

The FTDS was evaluated in three of the studies examined in this review, of which two defined a positive triage status as any positive criterion in the full scheme. In daily practice, only patients with any positive physiologic or anatomic criterion are advised to be treated in the highest level of trauma center available. This leads to an overestimation of sensitivity and underestimation of specificity. EMS professional judgment was the most applied field criterion, further emphasizing this assumption. It remains unclear how accurate the FTDS is in daily practice. Because EMS professional judgment is highly dependent on education level and experience, the protocol accuracy could be very different within and between regions.

The quality assessment of triage protocols requires large sample sizes. Subsequently, test accuracy was mostly evaluated retrospectively in the studies reviewed herein.[15-18] In these cases, interpretation of protocol criteria by EMS professionals was assumed to be identical to interpretation of retrospectively collected data by investigators. This simplification of reality will likely lead to biased diagnostic test accuracy and even more so when hospital data are used to replace missing pre-hospital data.

Three incomparable reference standards were used as surrogate markers for need of specialized trauma care. To date, it remains controversial which pediatric patients need expertise and resources of higher-level trauma centers and PTCs. Treating severely injured children (ISS ≥16) in higher-level trauma centers and PTCs has been shown to increase survival rates.[3,4] Major evidence for alternative reference standards is lacking. Children with an ISS of 25 or greater showed lower mortality rates at PTCs than at adult

or mixed trauma centers.[5] However, owing to the limited number and geographic distribution of PTCs, adult or mixed trauma centers provide care for the majority of children.[34]

In our opinion, an appropriate approach for investigating the test accuracy of triage protocols would be with a study population consisting of all children suspected of injury during field triage conducted by EMS professionals, independent of initial transport destination. The current literature often does not adhere to this approach. Differences in spectrum of disease, prevalence of patients requiring specialized trauma care, and patient characteristics lead to altered diagnostic performance. Thus, the results might not be representative for regional triage accuracy or protocol accuracy.[35] Undertriage and overtriage rates should be evaluated for complete trauma regions and EMS regions because knowing these rates is the key to improving pre-hospital trauma triage.

Strengths and Limitations

The strengths of this review include use of the latest methods for evidence searching and quality assessment. The review was limited to studies in a pre-hospital setting, discarding information from triage tools used in the emergency department or those used for prognostic purposes. This increased validity and clinical relevance of our findings for use in field triage.

The study results, although important, have several limitations. First, a lack of evidence exists on full diagnostic strategies in field triage of pediatric trauma patients. The isolated performance of a diagnostic test is difficult to interpret and could differ from a multicomponent context. Second, the included studies were of intermediate or low quality. Most studies retrospectively evaluated triage protocols not resembling daily practice. Third, the study populations were heterogeneous and the triage protocols evaluated by different reference standards were impossible to compare. In addition, all current evidence is from trauma systems in the United States, leaving it unclear whether our results would be transferable to trauma systems in other countries.

Conclusions

The goal of a field triage tool is to match the level of care needed by a patient with trauma to an acute care facility with the required amount of resources and expertise. The quality of the full diagnostic strategy used to transport the right patient to the right hospital is lacking. Current field triage tools misclassify a substantial number of injured children during field triage, potentially resulting in erroneous transportation destinations and preventable mortality. Increased efforts are needed to develop a highly sensitive and specific pediatric trauma triage tool to aid decision-making by EMS professionals.

REFERENCES

1. WHO. **World report on child injury prevention.** Geneva, 2008. https://www.who.int/violence_injury_prevention/child/injury/world_report/en/ (accessed Aug 20, 2017).

2. Amini R, Lavoie A, Moore L, Sirois MJ, Emond M. **Pediatric trauma mortality by type of designated hospital in a mature inclusive trauma system.** J Emerg Trauma Shock 2011; 4(1): 12-9.

3. Pracht EE, Tepas JJ, 3rd, Langland-Orban B, Simpson L, Pieper P, Flint LM. **Do pediatric patients with trauma in Florida have reduced mortality rates when treated in designated trauma centers?** J Pediatr Surg 2008; 43(1): 212-21.

4. Potoka DA, Schall LC, Gardner MJ, Stafford PW, Peitzman AB, Ford HR. I**mpact of pediatric trauma centers on mortality in a statewide system.** J Trauma 2000; 49(2): 237-45.

5. Sathya C, Alali AS, Wales PW, et al. **Mortality among injured children treated at different trauma center types.** JAMA Surg 2015; 150(9): 874-81.

6. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

7. Sampalis JS, Denis R, Frechette P, Brown R, Fleiszer D, Mulder D. **Direct transport to tertiary trauma centers versus transfer from lower level facilities: impact on mortality and morbidity among patients with major trauma.** J Trauma 1997; 43(2): 288-95; discussion 95-6.

8. Nirula R, Maier R, Moore E, Sperry J, Gentilello L. **Scoop and run to the trauma center or stay and play at the local hospital: hospital transfer's effect on mortality.** J Trauma 2010; 69(3): 595-9; discussion 9-601.

9. Leeflang MM, Bossuyt PM, Irwig L. **Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis.** J Clin Epidemiol 2009; 62(1): 5-12.

10. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld RA, Leenen LPH, van Heijl M. **Accuracy of prehospital triage protocols in selecting severely injured patients: a systematic review.** J Trauma Acute Care Surg 2017; 83(2): 328-39.

11. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** BMJ 2009; 339: b2535.

12. Bossuyt PM, Reitsma JB, Bruns DE, et al. **STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies.** BMJ 2015; 351: h5527.

13. Whiting PF, Rutjes AW, Westwood ME, et al. **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.** Ann Intern Med 2011; 155(8): 529-36.

14. Lerner EB, Drendel AL, Cushman JT, et al. **The ability of the physiologic criteria of the field triage guidelines to identify children who need the resources of a trauma center.** Academic Emergency Medicine 2016; 20(5): S61.

15. Newgard CD, Fu R, Zive D, et al. **Prospective validation of the national field triage guidelines for identifying seriously injured persons.** Journal of the American College of Surgeons 2016; 222(2): 146-58.e2.

16. Newgard CD, Zive D, Holmes JF, et al. **A multisite assessment of the American College of Surgeons Committee on trauma field triage decision scheme for identifying seriously injured children and adults.** Journal of the American College of Surgeons 2011; 213(6): 709-21.

17. Phillips S, Rond PC, 3rd, Kelly SM, Swartz PD. **The need for pediatric-specific triage criteria: results from the Florida Trauma**

Triage Study. Pediatric emergency care 1996; 12(6): 394-9.

18. Johnson WP. **Evaluation of the Pediatric Trauma Triage Checklist as a prehospital pediatric trauma triage tool for the state of Florida.** Prehospital and disaster medicine 1996; 11(1): 20-6.

19. MacKenzie EJ, Steinwachs DM, Ramzy AI. **Evaluating performance of statewide regionalized systems of trauma care.** J Trauma 1990; 30(6): 681-8.

20. Di Bartolomeo S, Tillati S, Valent F, Zanier L, Barbone F. **ISS mapped from ICD-9-CM by a novel freeware versus traditional coding: a comparative study.** Scand J Trauma Resusc Emerg Med 2010; 18: 17.

21. Tepas JJ, 3rd, Mollitt DL, Talbert JL, Bryant M. **The pediatric trauma score as a predictor of injury severity in the injured child.** J Pediatr Surg 1987; 22(1): 14-8.

22. Kaufmann CR, Maier RV, Rivara FP, Carrico CJ. **Evaluation of the Pediatric Trauma Score.** JAMA 1990; 263(1): 69-72.

23. Mackersie RC. **History of trauma field triage development and the American College of Surgeons criteria.** Prehosp Emerg Care 2006; 10(3): 287-94.

24. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Aug 20, 2017).

25. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

26. Janssens L, Holtslag HR, van Beeck EF, Leenen LP. **The effects of regionalization of pediatric trauma care in the Netherlands: a surveillance-based before-after study.** J Trauma Acute Care Surg 2012; 73(5): 1284-7.

27. Nakamura Y, Daya M, Bulger EM, et al. **Evaluating age in the field triage of injured persons.** Annals of emergency medicine 2012; 60(3): 335-45.

28. Peng J, Wheeler K, Groner JI, Haley KJ, Xiang H. **Undertriage of Pediatric Major Trauma Patients in the United States.** Clinical pediatrics 2017; 56(9): 845-53.

29. Xiang H, Wheeler KK, Groner JI, Shi J, Haley KJ. **Undertriage of major trauma patients in the US emergency departments.** The American journal of emergency medicine 2014; 32(9): 997-1004.

30. Corrado MM, Shi J, Wheeler KK, et al. **Emergency medical services (EMS) versus non-EMS transport among injured children in the United States.** Am J Emerg Med 2017; 35(3): 475-8.

31. Cicero MX, Cross KP. **Predictive value of initial Glasgow coma scale score in pediatric trauma patients.** Pediatric emergency care 2013; 29(1): 43-8.

32. Newgard CD, Rudser K, Atkins DL, et al. **The availability and use of out-of-hospital physiologic information to identify high-risk injured children in a multisite, population-based cohort.** Prehosp Emerg Care 2009; 13(4): 420-31.

33. Fleming S, Thompson M, Stevens R, et al. **Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies.** Lancet 2011; 377(9770): 1011-8.

34. Petrosyan M, Guner YS, Emami CN, Ford HR. **Disparities in the delivery of pediatric trauma care.** J Trauma 2009; 67(2 Suppl): S114-9.

35. Oudega R, Moons KG, Hoes AW. **Limited value of patient history and physical examination in diagnosing deep vein thrombosis in primary care.** Fam Pract 2005; 22(1): 86-91.

APPENDIX

No.	MEDLINE via OvidSP	Results
#1	((p?ediatric* or child* or teenager? or infant* or adolescen* or youth*).ti,ab. or exp child/ or exp adolescent/) and ((injur* or trauma? or wound*).ti,ab. or exp "Wounds and Injuries"/)	266485
#2	exp Emergency Medical Services/ or exp Emergency Service, Hospital/ or Rescue Work/ or exp Ambulances/ or (ambulance* or GEMS or HEMS or pre?hospital or out-of-hospital or field or ED).ti,ab.	685305
#3	(protocol? or flow?chart or decision scheme? or decision schema* or scoring* or score? or tool? or criteri* or triag* or priorit* or sort* or categoriz* or classif*).ti,ab. or exp Triage/	2312129
#4	(sensitivit* or specificit* or under?triage or over?triage or predictive value? or accurac* or roc or receiver operating characteristic?).ti,ab. or exp "Sensitivity and Specificity"/	1420242
#5	1 and 2 and 3 and 4	807

No.	PsycINFO via OvidSP	Results
#1	(p?ediatric* or child* or teenager? or infant* or adolescen* or youth*).ti,ab. and ((injur* or trauma? or wound*).ti,ab. or exp injuries/)	123314
#2	exp emergency services/ or (ambulance* or GEMS or HEMS or pre?hospital or out-of-hospital or field).ti,ab.	663633
#3	(protocol? or flow?chart or decision scheme? or decision schema* or scoring* or score? or tool? or criteri* or triag* or priorit* or sort* or categoriz* or classif*).ti,ab.	2741248
#4	exp test performance/ or exp test scores/ or exp test sensitivity/ or exp test specificity/ or (sensitivity* or specificit* or under?triage or over?triage or predictive value? or accuracy* or roc or receiver operating characteristic?).ti,ab.	1303288
#5	1 and 2 and 3 and 4	100

No.	Embase	Results
#1	('p?ediatric*' OR 'child*' OR 'teenager?' OR 'infant*' OR 'adolescen*' OR 'youth*' OR 'child'/exp OR 'adolescent'/exp OR 'juvenile'/de) AND ('injur*' OR 'trauma?' OR 'wound*' OR 'injury'/exp)	401840
#2	'pre-hospital':ab,ti OR 'out-of-hospital':ab,ti OR 'ambulance*' OR 'emergency medical service?' OR 'gems':ab,ti OR 'ems':ab,ti OR 'accident?' OR 'rescue*' OR 'emergency health service'/exp OR 'emergency medicine'/exp OR 'emergency'/exp OR 'paramedical personnel'/de OR 'rescue personnel'/exp	327215
#3	'protocol*' OR 'flow?chart?' OR 'scheme?' OR 'schema?' OR 'tool?' OR 'method*' OR 'system?' OR 'criteri*' OR 'priorit*' OR 'sort*' OR 'referr*' OR 'triag*'	15165827
#4	'sensitivit*' OR 'specificit*' OR 'under?triage' OR 'over?triage' OR 'predictive value?' OR 'accurac*' OR 'sensitivity and specificity'/exp OR 'predictive value'/exp OR 'diagnostic accuracy'/exp	1902743
#5	#1 and #2 and #3 and #4	778

No.	CENTRAL	Results
#1	('p?ediatric*':ab,ti or 'child*':ab,ti or 'teenager?':ab,ti or 'infant*':ab,ti or 'adolescen*':ab,ti or 'youth*':ab,ti or MeSH descriptor: [Child] explode all trees or MeSH descriptor: [Adolescent] explode all trees) and ('injur*':ab,ti or 'trauma?':ab,ti or 'wound*':ab,ti or MeSH descriptor: [Wounds and Injuries] explode all trees)	19748
#2	MeSH descriptor: [Emergency Medical Services] explode all trees or MeSH descriptor: [Emergency Service, Hospital] explode all trees or MeSH descriptor: [Rescue Work] explode all trees or MeSH descriptor: [Ambulances] explode all trees or 'ambulance*':ab,ti or 'GEMS':ab,ti or 'HEMS':ab,ti or 'pre?hospital':ab,ti or 'out-of-hospital':ab,ti or 'field':ab,ti or 'ED':ab,ti	22008
#3	'protocol?':ab,ti or 'flow?chart':ab,ti or 'decision scheme?':ab,ti or 'decision schema*':ab,ti or 'scoring*':ab,ti or 'score?':ab,ti or 'tool?':ab,ti or 'criteri*':ab,ti or 'triag*':ab,ti or 'priorit*':ab,ti or 'sort*':ab,ti or 'categoriz*':ab,ti or 'classif*':ab,ti or MeSH descriptor: [Triage] explode all trees	175298
#4	'sensitivit*':ab,ti or 'specificit*':ab,ti or 'under?triage':ab,ti or 'over?triage':ab,ti or 'predictive value?':ab,ti or 'accurac*':ab,ti or 'roc':ab,ti or 'receiver operating characteristic?':ab,ti or MeSH descriptor: [Sensitivity and Specificity] explode all trees	19079
#5	#1 and #2 and #3 and #4	19

CHAPTER III

Accuracy of pre-hospital trauma triage and field triage decision rules in children (P2-T2 study): an observational study

Lancet Child & Adolescent Health

Rogier van der Sluijs; Robin D. Lokerman; Job F. Waalwijk; Mariska A.C. de Jongh; Michael J.R. Edwards; Dennis den Hartog; Georgios F. Giannakópoulos; Pierre M. van Grunsven; Martijn Poeze; Luke P.H. Leenen; Mark van Heijl; for the Pre-hospital Trauma Triage Research Collaborative

ABSTRACT

Background

Adequate pre-hospital trauma triage is crucial to enable optimal care in inclusive trauma systems. Transport of children in need of specialized trauma care to lower-level trauma centers is associated with adverse patient outcomes. We aimed to evaluate the diagnostic accuracy of pediatric field triage based on patient destination and triage tools.

Methods

We did a multi-site observational study (P2-T2) of all children (aged <16 years) transported with high priority by ambulance from the scene of injury to any emergency department in seven of 11 inclusive trauma regions in the Netherlands. Diagnostic accuracy based on the initial transport destination was evaluated in terms of the undertriage rate (i.e., the proportion of patients in need of specialized trauma care who were initially transported to a lower-level pediatric or adult trauma center) and overtriage rate (i.e., the proportion of patients not requiring specialized trauma care who were transported to a level-I [highest level] pediatric trauma center). The Dutch National Protocol of Ambulance Services and Field Triage Decision Scheme triage protocols were externally validated using data from this cohort against an anatomical (Injury Severity Score [ISS] ≥16) and a resource-based reference standard.

Findings

Between Jan 1, 2015 and Dec 31, 2017, 12,915 children (median age 10.3 years, IQR 4.2 – 13.6) were transported to the emergency department with injuries. 4091 (31.7%) patients were admitted to hospital, of whom 129 (3.2%) patients had an ISS of 16 or greater and 227 (5.5%) patients used critical resources within a limited timeframe. Ten patients died within 24 h of arrival at the emergency department. Based on the primary reference standard (ISS ≥16), the undertriage rate was 16.3% (95%-CI, 10.8 – 23.7) and the overtriage rate was 21.2% (20.5 – 22.0). The National Protocol of Ambulance Services had a sensitivity of 53.5% (95%-CI, 43.9 – 62.9) and a specificity of 94.0% (93.4 – 94.6), and the Field Triage Decision Scheme had a sensitivity of 64.5% (54.1 – 74.1) and a specificity of 84.3% (83.1 – 85.5).

Interpretation

Too many children in need of specialized care were transported to lower-level pediatric or adult trauma centers, which is associated with increased mortality and morbidity. Current protocols cannot accurately discriminate between patients at low and high risk, and highly sensitive and child-specific triage tools need to be developed to ensure the right patient is transported to the right hospital.

Funding
The Netherlands Organization for Health Research and Development, Innovation Fund Health Insurers.

INTRODUCTION

Pediatric injuries are accountable for approximately 40% of all child deaths in developed countries worldwide.[1] Inclusive trauma systems were established to centralize patients, resources, and expertise to reduce mortality, lifelong disabilities, and costs. The higher the degree of centralization, the greater the consequences of inadequate field triage. Undertriage – transporting severely injured children to facilities without the required resources and expertise for optimal care (i.e., lower-level trauma centers) – is associated with higher mortality.[2,3] Conversely, overtriage – the transport of mildly injured children to higher-level pediatric trauma centers with a surplus of resources – results in excessive costs and exhaustive use of scarce resources.[4]

Field triage can be perceived as a three-step diagnostic strategy. First, Emergency Medical Service (EMS) professionals must determine a patient's resource-need at the scene of injury. This task can be challenging because of atypical injury presentations in children, limited time, and the few diagnostic modalities. Second, logistical constraints need to be considered, such as the capacity and proximity of trauma centers, and patient acuity. The third step is to determine the optimal transport destination in light of steps one and two.

Field triage tools can assist EMS professionals in the assessment of injury severity and subsequent resource-need. Allocation of injured children is guided by the National Protocol of Ambulance Services (NPAS) in the Netherlands.[5] This protocol is partly derived from the Field Triage Decision Scheme (FTDS), established by the American College of Surgeons Committee on Trauma.[6] The FTDS or similar combinations of physiologic, anatomic, and mechanism-related criteria are universally implemented in regionalized trauma systems across the world. When evaluated for a single inclusive trauma region in the Netherlands, the NPAS could not adequately select severely injured adults, with a sensitivity of only 36.2%.[7] Furthermore, a systematic review showed that few existing triage tools were child-specific and no single tool was able to attain an undertriage rate of less than 5%, as recommended by the American College of Surgeons Committee on Trauma.[6 8] Ultimately, no evidence was available on triage accuracy based on the transport destination of injured children: the essential question whether children requiring specialized care are in fact transported to higher-level pediatric trauma centers remained unanswered.

Accurate field triage is fundamental to properly functioning trauma systems. We designed the Pediatric Pre-hospital Trauma Triage (P2-T2) study to evaluate the quality of pediatric field triage in multiple EMSs and inclusive trauma regions in the Netherlands, based on protocol accuracy, protocol compliance, and destination-based mistriage rates. We also aimed to externally validate the NPAS and FTDS, two actively used field triage decision schemes.

METHODS

Study Design and Participants

The P2-T2 study was a multi-site, observational diagnostic study, which aimed to evaluate the quality of pediatric trauma field triage in the Netherlands.

All pediatric trauma patients (<16 years old) transported by EMSs with high priority were eligible for inclusion. Children transported by eight different EMSs (Amsterdam-Amstelland, Brabant Midden-West, Brabant-Noord, Gelderland-Zuid, Rotterdam-Rijnmond, Utrecht, Zaanstreek-Waterland, and Zuid-Holland Zuid) were included.

Patients not transported from the scene of injury to a trauma-receiving emergency department and patients transported to a trauma-receiving emergency department in one of the four non-participating trauma regions were excluded. Patients transported by helicopter were ineligible because of different triage strategies and protocols; few patients are transported by helicopter in the Netherlands because of the relatively short distances.

This study was reported in accordance with the Standards for Reporting of Diagnostic Accuracy Studies guidelines.[9] This study received approval from the Institutional Review Board of the University Medical Center Utrecht that the Medical Research Involving Human Subjects Act did not apply.

Procedures

The eight EMSs included in the study transport approximately 390,000 patients who are high priority annually, within an 8063 km^2 region, with a population of 6.5 million people. All EMSs in the Netherlands are provided by the government in partnership with private companies. Pre-hospital care systems are required to comply with the protocols of the Dutch Institute of Ambulance Care. All ground ambulances are staffed by a nurse, who is licensed to administer medical treatment at advanced life support level, and a dedicated driver. The initial transportation destination is controlled by the ambulance nurse. The participating EMSs are integrated into six regional inclusive trauma regions. Seven of 11

inclusive trauma regions in the Netherlands – encompassing six level-I pediatric trauma centers, seven level-I adult trauma centers, and 60 level-II or III pediatric or adult trauma centers – participated in this study. Each of the designated level-I (i.e., higher-level) pediatric trauma centers featured a pediatric intensive care unit (ICU) and offered trauma care at the highest level for severely injured children. Surrounding level-II and III facilities were considered lower-level trauma centers designated to treat mildly and moderately injured patients and did not feature pediatric ICUs. All participating EMSs and trauma regions are part of the Pre-hospital Trauma Triage Research Collaborative (PTTRC).

A novel tool (SelectAssist) was developed to aid trauma patient selection in the pre-hospital setting (see Appendix). Validation in a hold-out test dataset indicated an overall accuracy of 98.9% (95%-CI, 98.4 – 99.3). All initially selected patients were manually reviewed.

Pre-hospital electronic health records from participating EMSs were prospectively collected in a standardized manner from Jan 1, 2015 to Dec 31, 2017. All electronic health records were structured according to the template of the Dutch Basic Set of Ambulance Care and included patient demographics, vital signs, a description of the mechanism of injury, medical treatments administered, a primary survey, and a secondary survey.[10] Physiologic characteristics consisted of, among others, blood pressure, respiratory rate, and the Glasgow Coma Scale. Mechanisms of injury, suspected injuries, triage tool criteria, and relevant patient outcomes were collected from corresponding input fields. Higher-level trauma center criteria from both the NPAS and the FTDS were retrospectively applied to resemble strict adherence to the triage criteria. Free text fields available in the electronic health records were abstracted by research assistants and investigators to complement the collected data. If there was any uncertainty on the classification of a report, one different investigator was asked to review the patient record. If these investigators disagreed, a third investigator was asked to discuss the record until consensus was reached.

Pre-hospital data were linked with in-hospital data from the Dutch National Trauma Registry to construct the final dataset.[10] This registry is nationwide covering all trauma-related hospital admissions from every emergency department in the Netherlands. The registry includes relevant patient outcomes, such as: Injury Severity Score (ISS), mortality, ICU admission, and early critical-resource use.[11] The Abbreviated Injury Scale version 2005, update 2008, was used by trained trauma data managers to calculate the ISS after the final diagnosis was made.[12] Patients discharged at the emergency department to their home environment were considered not to be severely injured, nor to have used any critical resource within a predefined time frame. These assumptions were verified and confirmed in hand-collected data from a previous study (n=4950).[7]

A combined deterministic and probabilistic linkage scheme was used to match pre-hospital patient records to patient outcomes. A novel tool (LinkAssist; see Appendix) was developed for probabilistic record linkage of pre-hospital and in-hospital data. The record linkage strategy was validated to be 100.0% (95%-CI, 100.0 – 100.0) accurate in previously unseen data. All patients requiring specialized trauma care admitted to any of the participating emergency departments were checked by hand.

The NPAS was used in daily practice by all EMSs in the Netherlands to select children in need of specialized trauma care during field triage. The NPAS (Triage Choice of Hospital, version 8.1) advised transport to a level-I pediatric trauma center when at least one physiologic or injury-related criterion was met. The FTDS (2011 version) was the only actively used and evaluated protocol identified in our recent systematic review on the accuracy of pediatric field triage.[8] The FTDS advised to transport a patient to the highest level of care within a trauma system when at least one physiologic or injury-related criterion was met.

Outcomes

The primary outcome was the diagnostic accuracy (undertriage and overtriage rates) of the full triage strategy on the basis of the initial transport destination. Patients requiring specialized trauma care transported to level-I pediatric trauma centers and patients not in need of specialized trauma care transported to lower-level trauma centers were considered adequately triaged. The undertriage rate was defined as the proportion of patients in need of specialized trauma care who were initially transported to a lower-level pediatric or adult trauma center, while the overtriage rate was defined as the proportion of patients not requiring specialized trauma care who were transported to a level-I pediatric trauma center.

The primary reference standard for need of specialized care was defined as an ISS of 16 or greater (possible range 0–75), as is suggested by the American College of Surgeons Committee on Trauma to evaluate triage accuracy.[6] As the ISS is based on anatomic criteria, it is assumed to be consistent with the patient status on-scene and was therefore used as a diagnostic reference standard. A secondary, resource-based, reference standard was adopted that was targeted on early critical-resource use. Early critical-resource use was defined as a composite endpoint consisting of intubation in the pre-hospital setting, ICU admission after discharge from the emergency department, major surgical intervention within 12 h, major radiological intervention within 12 h, or death within 24 h after arrival the emergency department (Appendix). This definition is similar to previous research with resource-based reference standards.[13,14]

Table 1 NPAS and FTDS Triage Tools

	Triage tools	
	NPAS (Choice of hospital, version 8.1)	**FTDS (2011 version)**
Physiologic characteristics*	Glasgow Coma Scale <9 or deteriorating; airway, breathing, or circulation cannot be stabilized; Revised Trauma Score <11 or Pediatric Trauma Score <9; anisocoria; hypothermia ≤32°C	Glasgow Coma Scale <14; systolic blood pressure (mm Hg) <90 mm Hg; respiratory rate <10 or >29 per min or <20 in infant aged <1 year
Anatomy of suspected injuries†	All penetrating injuries to the head, thorax, and abdomen; flail chest; two or more long-bone fractures; amputation proximal to wrist and ankle; unstable pelvic fracture; paralysis	All penetrating injuries to head, neck, torso, and extremities proximal to elbow and knee; flail chest; two or more proximal long-bone fractures; amputation proximal to wrist and ankle; pelvic fractures; paralysis; open or depressed skull fracture; crushed, degloved, or mangled extremity
Mechanism of injury‡	Falls: >5 m or 3 times the height of the patient; high-risk motor vehicle crash; deformity >50 cm or intrusion >30cm occupant site; ejection from automobile; death in same vehicle; more than 65 km/h; pedestrian–automobile collision >10 km/h	Falls: >6 m, or >3 m or 2–3 times the height of the child for children <15 years old; high-risk auto crash; intrusion: >30 cm occupant site; >46 cm any site; ejection (partial or complete) from automobile; death in same passenger compartment; motorcycle crash >32 km/h, pedestrian or cyclist–automobile collision where pedestrian or cyclist was thrown, run over, or hit with substantial (>32 km/h) impact; vehicle telemetry data consistent with high risk of injury

The order of variables of the NPAS and the FTDS were changed to facilitate comparison between the two triage tools. Units in the FTDS were converted to SI units. Special patient or system considerations were removed (e.g., pregnancy). Abbreviations: FTDS, Field Triage Decision Scheme; NPAS, *National Protocol of Ambulance Services*. *Patients fulfilling any of the physiologic criteria should be transported to a level-I pediatric trauma center or the highest level of care within the trauma system in case of the FTDS. †Patients suspected of any of these injury characteristics should be transported to a level-I pediatric trauma center (NPAS) or the highest level of care within the trauma system (FTDS). ‡Patients should be transported to a regional trauma center, but not necessarily the highest level of care within the system.

The secondary outcomes were diagnostic accuracy (sensitivity, specificity, negative predictive value, positive predictive value, positive likelihood ratio, and negative likelihood ratio) of the NPAS and the FTDS triage tools (Table 1), and the compliance of EMS professionals to the NPAS. Sensitivity was defined as the proportion of patients in need of specialized trauma care (ISS ≥16 or early critical-resource use), who were correctly identified by the triage tool, regardless of the destination facility. Specificity was defined as the proportion of children not in need of specialized care, which was accurately classified by the triage tool, again disregarding the destination facility. Triage tool compliance by EMS professionals was assessed by quantifying the discrepancy between the trauma center level suggested by the triage tool and the designated level of the actual transport destination.

Statistical Analysis

Descriptive statistics were calculated using frequencies and percentages for categorical variables, and median values with IQRs for continuous variables. Contingency tables were constructed for each pair of index tests and reference standards. Diagnostic accuracy metrics were derived from each contingency table. Binomial Agresti-Coull 95%-CIs were calculated when applicable.[15] The log-method was used to define 95%-CIs for ratios.[16] Patient height was estimated on the basis on age, sex, and mean values per age group in the Dutch population. The haversine method was used to calculate the great-circle-distance (i.e., as the crow flies) between the scene of injury and surrounding hospitals. The sample size was determined on the basis of prior research in an adult population.[7]

Multiple imputation was used to address missing values. Variables present in the triage criteria with missing values were Glasgow Coma Scale, systolic blood pressure, and respiratory rate. Missing values were imputed on the basis of a predictor matrix that included, among others, vital signs measured in the emergency department, age, sex, and patient outcomes of both children and adults. Multi-level multiple imputation methods that account for clustering across sites were adopted using the R-package micemd.[17] 48 datasets were generated, based on 20 iterations per set. Analyses were applied to each of the 48 datasets. Results for all analyses were averaged to calculate point estimates. Confidence intervals were calculated in accordance with Rubin's rules.[18]

Role of the Funding Source

The funders of the study had no role in the study design, data collection, analyses, interpretation of the data, or the writing of the final report. RvdS, JFW, RDL had full access to all the data and the corresponding author had the final responsibility to submit for publication.

RESULTS

Between Jan 1, 2015 and Dec 31, 2017, approximately 1.5 million patient records were identified of which 631,475 patients were transported to an emergency department with high priority (Figure). 12,931 of these were pediatric trauma patients eligible for inclusion, and after excluding 16 patients who were lost to follow-up, 12,915 patients were included in the analysis. Pre-hospital variables with missing values were systolic blood pressure (49,411 [29.9%] of 165,404), Glasgow Coma Scale (24,169 [14.6%]), and respiratory rate (60,899 [36.8%]).

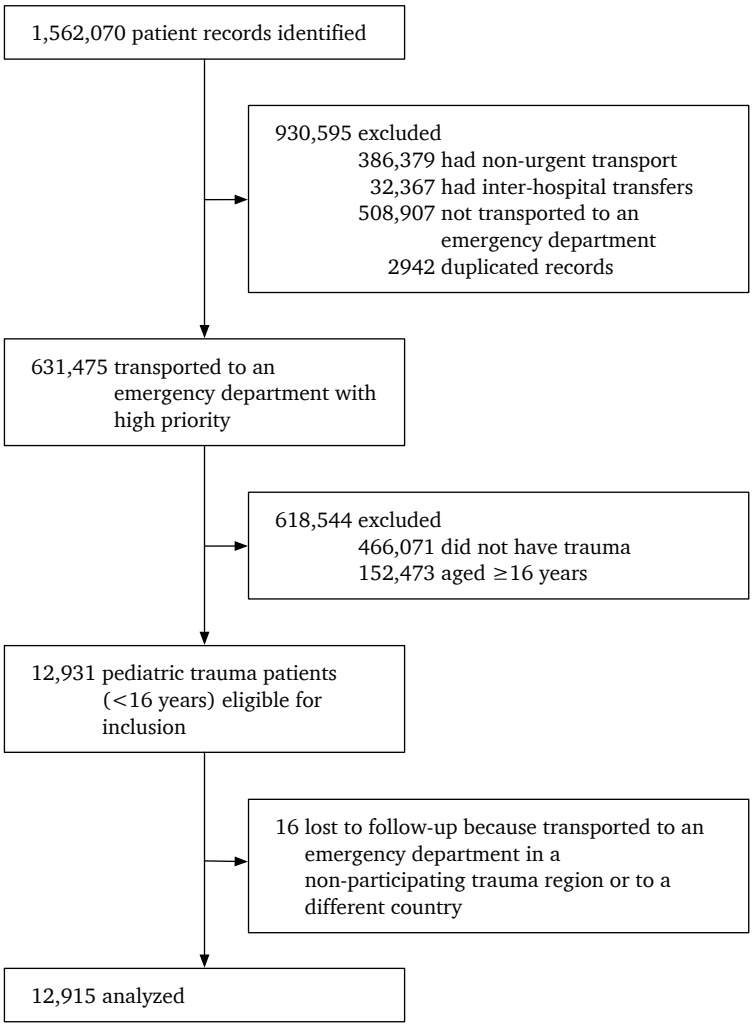


Figure Study Profile

The median age of the included patients was 10.3 years (IQR 4.2 – 13.6), 7503 (58.1%) were boys, and 5412 (41.9%) were girls (Table 2). The median systolic blood pressure increased per age group from 106 mm Hg for children aged 0–1 years to 124 mm Hg for those aged 11–15 years, whereas median respiratory rates and heart rates decreased with increasing age. The median distance to a higher-level trauma center was 11.3 km (IQR, 5.3 – 26.7). The closest trauma center was bypassed 1314 times (10.2%). The most common injury patterns were head injuries (3683 [28.5%] of 12,915) and injuries in the extremities (3780 [29.3%]). Injury patterns differed considerably between age groups. Head injuries were

Table 2 Patients Characteristics in the Cohort per Age Group

	Age groups				
	All patients	**0 to <2 years**	**2 to <6 years**	**6 to <11 years**	**11 to <16 years**
Demographic characteristics					
Age, years	10.3 (4.2 – 13.6)	1.1 (0.7 – 1.5)	3.7 (2.8 – 4.9)	8.8 (7.4 – 10.1)	13.8 (12.6 – 14.9)
Boys	7503 (58.1%)	936 (54.7%)	1524 (61.5%)	1645 (58.8%)	3398 (57.3%)
Girls	5412 (41.9%)	775 (45.3%)	953 (38.5%)	1155 (41.2%)	2529 (42.7%)
Pre-hospital physiologic characteristics					
Systolic blood pressure, mm Hg	120 (110 – 130)	106 (92 – 126)	110 (100 – 120)	114 (105 – 124)	124 (114 – 134)
Systolic blood pressure <90 mm Hg	488 (3.8%)	93 (5.4%)	140 (5.6%)	108 (3.8%)	148 (2.5%)
Respiratory rate per min	18 (15 – 20)	24 (16 – 30)	20 (16 – 24)	18 (16 – 20)	16 (14 – 20)
Respiratory rate <10 or >29 per min	1046 (8.1%)	521 (30.5%)	237 (9.6%)	102 (3.7%)	186 (3.1%)
Heart rate, bpm	93 (80 – 110)	120 (98 – 134)	103 (90 – 120)	92 (82 – 103)	88 (78 – 100)
Oxygen saturation, %	99 (98 – 99)	99 (97 – 100)	99 (98 – 99)	99 (98 – 99)	99 (98 – 99)
Glasgow Coma Scale score	15 (15 – 15)	15 (15 – 15)	15 (15 – 15)	15 (15 – 15)	15 (15 – 15)
Characteristics of suspected injuries					
Head	3683 (28.5%)	779 (45.5%)	1019 (41.1%)	660 (23.6%)	1225 (20.7%)
Thorax	218 (1.7%)	4 (0.2%)	19 (0.8%)	59 (2.1%)	136 (2.3%)
Abdomen	263 (2%)	6 (0.4%)	34 (1.4%)	87 (3.1%)	136 (2.3%)
Extremities	3780 (29.3%)	55 (3.2%)	370 (14.9%)	908 (32.4%)	2447 (41.3%)
Mechanism of injury					
Falls	6492 (50.3%)	1122 (65.6%)	1508 (60.9%)	1505 (53.8%)	2357 (39.8%)
>3 m	140 (1.1%)	10 (0.6%)	35 (1.4%)	48 (1.7%)	47 (0.8%)
≥3 times the child's body length	218 (1.7%)	152 (8.9%)	34 (1.4%)	22 (0.8%)	10 (0.2%)
Fall from stairs	1042 (8.1%)	349 (20.4%)	441 (17.8%)	114 (4.1%)	138 (2.3%)
Motor vehicle incident	2023 (15.7%)	87 (5.1%)	297 (12%)	470 (16.8%)	1169 (19.7%)
Bicycle incident	2464 (19.1%)	74 (4.3%)	308 (12.4%)	466 (16.6%)	1616 (27.3%)
Pedestrian–automobile collision	399 (3.1%)	11 (0.6%)	110 (4.4%)	157 (5.6%)	121 (2%)
Bicycle–automobile collision	773 (6%)	7 (0.4%)	42 (1.7%)	117 (4.2%)	607 (10.2%)

Table 2 Continued

	Age groups				
	All patients	**0 to <2 years**	**2 to <6 years**	**6 to <11 years**	**11 to <16 years**
Burns	478 (3.7%)	218 (12.7%)	100 (4%)	72 (2.6%)	88 (1.5%)
Suffocation	133 (1%)	71 (4.1%)	45 (1.8%)	6 (0.2%)	11 (0.2%)
Submersion	75 (0.6%)	17 (1%)	34 (1.4%)	10 (0.4%)	14 (0.2%)
Sport related	2688 (20.8%)	7 (0.4%)	85 (3.4%)	592 (21.1%)	2004 (33.8%)
Violence	214 (1.7%)	6 (0.4%)	10 (0.4%)	19 (0.7%)	179 (3%)
Outcomes					
In–hospital stay	4091 (31.7%)	893 (52.2%)	1050 (42.4%)	901 (32.2%)	1247 (21%)
ISS*	2 (1 – 4)	2 (1 – 2)	2 (1 – 4)	4 (2 – 5)	4 (2 – 5)
ISS ≥16*	129 (3.2%)	21 (2.4%)	27 (2.6%)	29 (3.2%)	52 (4.2%)
Abbreviated Injury Scale score ≥3, per ISS region*					
Head and neck	202 (4.9%)	36 (4.0%)	45 (4.3%)	44 (4.9%)	77 (6.2%)
Face	4 (<0.1%)	1 (0.1%)	0	2 (0.2%)	1 (<0.1%)
Thorax	71 (1.7%)	10 (1.1%)	17 (1.6%)	15 (1.7%)	29 (2.3%)
Abdomen	39 (1.0%)	1 (0.1%)	3 (0.3%)	10 (1.1%)	25 (2%)
Extremities	225 (5.5%)	10 (1.1%)	38 (3.6%)	75 (8.3%)	102 (8.2%)
External	56 (1.4%)	19 (2.1%)	25 (2.4%)	7 (0.8%)	5 (0.4%)
Early critical–resource use*	227 (5.5%)	38 (4.3%)	55 (5.2%)	52 (5.8%)	82 (6.6%)
Discharge from emergency department to intensive care unit	155 (3.8%)	31 (3.5%)	43 (4.1%)	34 (3.8%)	47 (3.8%)
Out–of–hospital intubation	90 (2.2%)	13 (1.5%)	22 (2.1%)	14 (1.6%)	41 (3.3%)
Major interventions <12 h	48 (1.2%)	4 (0.4%)	5 (0.5%)	19 (2.1%)	20 (1.6%)
Mortality <24 h	10 (0.2%)	1 (0.1%)	5 (0.5%)	2 (0.2%)	2 (0.2%)
Transportation to level–I pediatric trauma center	2823 (21.9%)	515 (30.1%)	700 (28.3%)	647 (23.1%)	961 (16.2%)
Transportation to non–pediatric or lower–level trauma center	10,092 (78.1%)	1196 (69.9%)	1777 (71.7%)	2153 (76.9%)	4966 (83.8%)

Data are median (IQR) or n (%). Values derived from multiply imputed variables were rounded to zero decimals. Abbreviations: ISS, *Injury Severity Score*. *Hospitalized patients only.

Table 3 Presence of NPAS and FTDS Criteria in the Cohort

	Number of patients (n=12,915)
NPAS and FTDS	
Flail chest	1 (<0.1%)
Amputation proximal to wrist and ankle	1 (<0.1%)
Paralysis	14 (0.1%)
NPAS	
Glasgow Coma Scale score <9 or deteriorating	171 (1.3%)
Revised Trauma Score <11 or Pediatric Trauma Score <9	490 (3.8%)
Anisocoria	5 (<0.1%)
Hypothermia ≤32 °C	9 (0.1%)
All penetrating injuries to the head, thorax, and abdomen	19 (0.1%)
Two or more long-bone fractures (including tibial fractures)	8 (0.1%)
Unstable pelvic fracture	7 (0.1%)
Fall >5 meter or three times the height of the patient	227 (5.8%)
Motor vehicle deformity >50 cm or intrusion >30 cm occupant site	12 (0.1%)
Motor vehicle ejection	3 (<0.1%)
Motor vehicle incident with death in same vehicle	2 (<0.1%)
Motor vehicle incident >65 km/h	153 (1.2%)
Pedestrian–automobile collision >10 km/h	395 (3.1%)
FTDS	
Glasgow Coma Scale score <14	693 (5.4%)
Systolic blood pressure (mm Hg) <90 mm Hg	488 (3.8%)
Respiratory rate <10 or >29 per min or <20 in infants aged <1 year	1001 (7.7%)
All penetrating injuries to head, neck, torso, and extremities proximal to elbow and knee	24 (0.2%)
Two or more proximal long-bone fractures	4 (<0.1%)
Pelvic fractures	50 (0.4%)
Open or depressed skull fracture	20 (0.2%)
Crushed, degloved, or mangled extremity	13 (0.1%)

Table 3 Continued

	Number of patients (n=12,915)
Fall >6 meter, or >3 meter or 2–3 times the height of the child for children <15 years old	614 (15.6%)
Motor vehicle intrusion: >30 cm occupant site; >46 cm any site	12 (0.1%)
Motor vehicle ejection (partial or complete)	3 (<0.1%)
Motor vehicle incident with death in same passenger compartment	0
Motorcycle crash >32 km/h	50 (0.4%)
Pedestrian or bicyclist–automobile collision, thrown, run over, or with significant (>32 km/h) impact	477 (3.7%)
Vehicle telemetry data consistent with high risk of injury	0

Values are presented as n (%). Abbreviations: FTDS, *Field Triage Decision Scheme (version of 2011)*; NPAS, *National Protocol of Ambulance Services (Choice of hospital, version 8.1)*.

more prevalent in children aged up to six years, whereas those aged 6–15 years had injured extremities more often. More than half of the patients had a fall (6492 [50.3%]), 2464 (19.1%) had a bicycle incident, 2023 (15.7%) were involved in a motor vehicle incident, and 478 (3.7%) had burn injuries. 2688 (20.8%) patients had a sport-related injury.

4,091 (31.7%) patients were hospitalized after evaluation at the emergency department. The median ISS of hospitalized patients was 2 (IQR, 1 – 4). 129 (3.2%) patients who were hospitalized with an ISS of 16 or greater were considered in need of specialized trauma care based on the primary reference standard. 202 (4.9%) of 4091 patients who were hospitalized had an Abbreviated Injury Scale of 3 or greater in the head or neck. Likewise, 225 (5.5%) had a score of 3 or greater in the extremities. Early critical-resources were used by 227 (5.5%) of the hospitalized patients, of which 90 (2.2%) were intubated before arriving at the emergency department, and 48 (1.2%) underwent a major surgical intervention (e.g., craniotomy, intracranial pressure monitoring, damage control thoracotomy/laparotomy) or radiological intervention. Ten patients (0.1%) died within 24 h after arriving at the emergency department, all of whom had impaired vital signs and the median Glasgow Coma Scale was 3 (IQR, 3 – 7). Seven of these patients had submersion injuries, two patients were involved in motor vehicle accidents, and one patient died as a result of an explosion. The percentage of children transported to a level-I pediatric trauma center decreased from 30.1% to 16.2% with older age groups. The individual NPAS and FTDS criteria observed in the current cohort are shown in

Table 4 Diagnostic Accuracy of Pediatric Field Triage based on Destination, the NPAS, and the FTDS

	Injury Severity Score ≥16	**Early critical-resource use**
Destination-based		
Transported to level-I PTC		
With outcome (TP)	108	194
Without outcome (FP)	2715	2629
Transported to lower-level PTC or adult trauma center		
With outcome (FN)	21	33
Without outcome (TN)	10,071	10,059
Undertriage	16.3 (10.8 – 23.7)	14.5 (10.5 – 19.7)
Overtriage	21.2 (20.5 – 22.0)	20.7 (20.0 – 21.4)
NPAS*		
Positive on criteria		
With outcome (TP)	69.0	114.5
Without outcome (FP)	765.7	720.2
Negative on criteria		
With outcome (FN)	60.0	112.5
Without outcome (TN)	12,020.3	11,967.8
Sensitivity	53.5 (43.9 – 62.9)	50.4 (43.6 – 57.3)
Specificity	94.0 (93.4 – 94.6)	94.3 (93.7 – 94.9)
Positive predictive value	8.3 (6.5 – 10.4)	13.7 (11.4 – 16.4)
Negative predictive value	99.5 (99.4 – 99.6)	99.1 (98.9 – 99.2)
Positive likelihood ratio	8.9 (4.4 – 18.2)	8.9 (4.9 – 16.2)
Negative likelihood ratio	0.5 (0.4 – 0.6)	0.5 (0.5 – 0.6)
FTDS†		
Positive on criteria		
With outcome (n)	83.2	137.3
Without outcome (n)	2006.3	1952.3
Negative on criteria		
With outcome (n)	45.8	89.7

Table 4 Continued

	Injury Severity Score ≥16	**Early critical-resource use**
Without outcome (n)	10,779.7	10,735.7
Sensitivity	64.5 (54.1 – 74.1)	60.5 (52.8 – 67.8)
Specificity	84.3 (83.1 – 85.5)	84.6 (83.4 – 85.8)
Positive predictive value	4.0 (3.2 – 5.0)	6.6 (5.5 – 7.8)
Negative predictive value	99.5 (99.4 – 99.6)	99.2 (99.0 – 99.3)
Positive likelihood ratio	4.1 (2.8 – 6.1)	3.9 (3.0 – 5.1)
Negative likelihood ratio	0.4 (0.3 – 0.5)	0.5 (0.4 – 0.6)

Data are n; 95% CIs are given in parentheses were applicable. Abbreviations: FN, *false negative*; FP, *false positive*; FTDS, *Field Triage Decision Scheme*; NPAS, *National Protocol of Ambulance Services*; PTC, *Pediatric Trauma Center*; TN, *true negative*; TP, *true positive*. *Choice of hospital, version 8.1. †Version of 2011.

Table 3. One or more of the NPAS physiologic criteria were met by 799 (6.2%) children. By contrast, 1993 (15.4%) children fulfilled one or more of FTDS physiologic criteria. Anatomic criteria of the NPAS were met by 47 (0.4%) and anatomic criteria of the FTDS were met by 121 (0.9%) children. The NPAS was considered positive in 835 (6.5%) patients, whereas the FTDS generated a positive test result in 2090 (16.2%) children.

The undertriage rate in the current cohort was 16.3% (95%-CI, 10.8 – 23.7) based on the primary reference standard (ISS ≥16), while the overtriage rate was 21.2% (95%-CI, 20.5 – 22.0; Table 4). Evaluation using the secondary reference standard (early critical-resource use) resulted in an undertriage rate of 14.5% (9%-CI, 10.5 – 19.7) and an overtriage rate of 20.7% (95%-CI, 20.0 – 21.4). 50 (26.7%; 95%-CI, 20.9 – 33.5) patients with severe traumatic brain injury (Abbreviated Injury Scale score ≥3) were first transported to lower-level trauma centers. Eight of ten patients who died were initially transported to a level-I pediatric trauma center. The median distance from the scene of injury of patients who were undertriaged to a higher-level trauma center was 22.4 km (IQR, 9.0 – 31.6) with a maximum of 51.0 km.

The sensitivity of the NPAS was 53.5% (95%-CI, 43.9 – 62.9) and the specificity was 94.0% (95%-CI, 93.4 – 94.6), based on the primary reference standard. The FTDS was 64.5% (95%-CI, 54.1 – 74.1) sensitive to select children with an ISS of 16 or greater and had a specificity of 84.3% (95%-CI, 83.1 – 85.5). Evaluation of the NPAS using the secondary, resource-based, reference standard resulted in a sensitivity of 50.4% (95%-CI, 43.6 – 57.3) and a specificity of 94.3% (95%-CI, 93.7 – 94.9). Based on the secondary

reference standard, the FTDS had a sensitivity of 60.5% (95%-CI, 52.8 – 67.8) and a specificity of 84.6% (95%-CI, 83.4 – 85.8). EMS professionals complied to a positive result of the NPAS in 44.7% (95%-CI, 40.9 – 48.5) of the cases. Compliance to a negative advice was 79.7% (95%-CI, 79.0 – 80.4), although the NPAS does not obligate EMS professionals to transport patients who are mildly injured to lower-level trauma centers.

DISCUSSION

This multi-site study investigated the accuracy of pediatric pre-hospital trauma triage based on patient transport destination, triage protocols, and protocol compliance. For the primary outcome, we found that 16% of children requiring specialized trauma care were not transported to a level-I pediatric trauma center. The full triage strategy was therefore unable to attain a satisfactory undertriage rate of less than 5%, which might have led to avoidable adverse patient outcomes. Moreover, the NPAS and FTDS triage protocols were externally validated in the current cohort and both were insensitive and moderately-to-highly specific using an anatomical and a resource-based reference standard. Triage protocols, however, must have very high sensitivities, indicating that children designated as low risk do not include patients in need of specialized trauma care, to avoid undertriage. Both protocols had high negative predictive value and low positive predictive value, which was likely influenced by the low prevalence of a positive reference standard.[19] Conformity to the NPAS in daily practice was low, but did lead to lower undertriage rates as compared with strict protocol adherence.

The strengths of this study are its generalizability, the robust and sensitive methods used to select patients, the record linkage strategy, and the methodology used to construct the cohort. First, we selected EMSs with urban, suburban, and rural service areas. These heterogeneous EMSs were chosen to increase the generalizability of the results. The effect of inadequate triage tools is likely applicable to other inclusive trauma systems with similar (dichotomous) triage strategies, whereas destination-based accuracy might be different because of geographical differences, local regulations, and dissimilarities in the education of EMS professionals. These results might also extrapolate to trauma team activation protocols based on similar criteria in both inclusive trauma systems and exclusive trauma centers, although this possibility needs to be verified in separate studies.

Second, children were not solely selected based on the chief complaint (e.g., traumatology), but all children transported by EMSs to emergency departments in participating trauma regions were thoroughly screened in order not to miss any patient who needed trauma-related emergency department evaluation. Only 16 patients were lost to follow-up after transport to a non-participating trauma center.

Third, all emergency departments in the participating trauma regions contributed to data collection, which enabled the most appropriate type of study population to investigate triage rates and to validate triage protocols. Another advantage of this study was the standardized data collection, and the calculation of ISSs based on the Abbreviated Injury Scale. This approach is superior to ICD-9 derived ISSs.[20]

To interpret our results, it is important to note that the study was not statistically powered to calculate triage rates in different subgroups (e.g., per resource, age group, or region) because of the low prevalence of children in need of specialized trauma care. These calculations would have been insightful, since injury patterns differ between age groups, and the density, proximity, and capacity of pediatric trauma centers differ between trauma regions.

Both an anatomical and a resource-based reference standard were used to evaluate the diagnostic accuracy of the NPAS and the FTDS. Trauma systems are traditionally evaluated using the ISS, but resource-based reference standards are increasingly being proposed as a supposedly better alternative to determine need of specialized trauma care.[13,21] It could be argued that resource use directly anticipates on the consequences of trauma care regionalization (e.g., centralization of resources), whereas ISS merely acts as a surrogate marker. However, the downside of a resource-based reference standard, is the fact that it is a composite endpoint dependent on the current triage strategy (i.e., resources might be unavailable at the transport destination and thus not used), whereas the ISS is not.

The high undertriage rate might be partly attributable to the low sensitivity of the NPAS, although it was higher in comparison to a more select adult population.[7] The diagnostic accuracy of the FTDS was previously evaluated in four studies done in the United States. Lerner and colleagues evaluated the accuracy of the physiologic step and the complete FTDS (the entire triage algorithm, including mechanism of injury, and special considerations) in a prospective study in three pediatric trauma centers based on a resource-based reference standard.[13,22] The reported sensitivities were 49% for the physiologic step of the FTDS and 65% for the complete FTDS. These results cannot be directly compared to the current study, since they were derived from a selected population of children admitted to pediatric trauma centers and focused on different steps of the FTDS.[13,22] Two studies by Newgard and colleagues evaluated the complete FTDS (version 2006) based on an ISS of 16 or greater.[20,23] Sensitivity ranged between 84% and 87%. Judgment by EMS professionals was the most frequently used triage criterion. This suggests that the triage criteria were unable to capture the heterogeneous population in need of specialized trauma care. Our study evaluated the physiologic and anatomic criteria, which makes it unsuitable for comparison. Finally, one study evaluated the compliance of EMS professionals to the physiologic step of the FTDS using transport

destination as the outcome.[24] It showed that more than one-quarter of the children positively identified by the FTDS physiologic criteria were transported to a lower-level trauma center. Our study reports even higher non-compliance to the NPAS.

The poor performance of the evaluated field triage protocols is likely to be multicausal. Several scenarios, mechanisms of injury, and heterogeneous injury patterns can lead to severe injuries. Criteria in current decision schemes are generic, do not interact with other criteria, and are therefore unable to produce advice on a patient level. A probability or risk score might be a better alternative, since triage protocols are only one component of the full triage strategy.[14,25] Furthermore, dichotomization of physiologic criteria leads to a loss of information and the current cutoff points are not child-specific (see Appendix).[26]

Further research could focus on child-specific, highly discriminative predictors for need of specialized trauma care. New prediction models could be developed and validated in order to aid EMS professionals during the challenging process of field triage.

Undertriage and overtriage rates are key metrics to evaluate the functioning of inclusive trauma systems. This study reveals undertriage rates that need improvement to enable maximally cost-efficient care. After all, centralization of resources and expertise (i.e., regionalized trauma care) could be detrimental if patients are not transported to the right hospital.

Declaration of interests
We declare no competing interests.

Acknowledgements
This study was partly funded by grants from the Netherlands Organization for Health Research and Development (ZonMw) and the Innovation Fund Health Insurers. We thank the research staff from all participating sites and all participating members of the Pre-hospital Trauma Triage Research Collaborative. We thank our research assistants: Alexander de la Mar, Annemiek Vuurens, Dunja Scheepmaker, and Toril Lintzen. Finally, we would like to thank Jill Whittaker for her advice on writing style.

Author contributions
RvdS, MP, LPHL, and MvH wrote the statistical analysis plan, analyzed the data, and drafted the manuscript. All authors contributed to the study design, interpreted the data, and revised the manuscript.

REFERENCES

1. WHO. **World report on child injury prevention.** Geneva, 2008. https://www.who.int/violence_injury_prevention/child/injury/world_report/en/ (accessed Aug 20, 2019).

2. Potoka DA, Schall LC, Gardner MJ, Stafford PW, Peitzman AB, Ford HR. **Impact of pediatric trauma centers on mortality in a statewide system.** J Trauma 2000; 49(2): 237-45.

3. Sathya C, Alali AS, Wales PW, et al. **Mortality among injured children treated at different trauma center types.** JAMA Surg 2015; 150(9): 874-81.

4. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

5. Ambulancezorg Nederland**. Landelijk Protocol Ambulancezorg (versie 8.1).** Zwolle, 2016. https://www.ambulancezorg.nl/themas/kwaliteit-van- zorg/protocollen-en-richtlijnen/landelijk-protocol-ambulancezorg (accessed Aug 20, 2019).

6. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Aug 20, 2019).

7. Voskens FJ, van Rein EAJ, van der Sluijs R, et al. **Accuracy of prehospital triage in selecting severely injured trauma patients.** JAMA Surg 2018; 153(4): 322-7.

8. van der Sluijs R, van Rein EAJ, Wijnand JGJ, Leenen LPH, van Heijl M. **Accuracy of pediatric trauma field triage: a systematic review.** JAMA Surg 2018; 153(7): 671-6.

9. Bossuyt PM, Reitsma JB, Bruns DE, et al. **STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies.** BMJ 2015; 351: h5527.

10. van Rein EAJ, Jochems D, Lokerman RD, et al. **Diagnostic value of emergency medical services provider judgement in the identification of head injuries among trauma patients.** Eur J Neurol 2019; 26(2): 274-80.

11. 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(3): 187-96.

12. Gennarelli T, Wodzin E. **The Abbreviated Injury Scale 2005: update 2008**. Association for the Advancement of Automotive Medicine; Des Plaines, IL: 2008.

13. Lerner EB, Cushman JT, Drendel AL, et al. **Effect of the 2011 Revisions to the field triage guidelines on under- and over-triage rates for pediatric trauma patients.** Prehosp Emerg Care 2017; 21(4): 456-60.

14. Van der Sluijs R, Debray TPA, Poeze M, Leenen LPH, van Heijl M. **Development and validation of a novel prediction model to identify patients in need of specialized trauma care during field triage: design and rationale of the GOAT study.** Diagnostic and Prognostic Research 2019; 3(12).

15. Agresti A, Coull BA. **Approximate is better than "exact" for interval estimation of binomial proportions.** Am Stat 1998; 52(2): 119-26.

16. Simel DL, Samsa GP, Matchar DB. **Likelihood ratios with confidence: sample size estimation for diagnostic test studies.** J Clin Epidemiol 1991; 44(8): 763-70.

17. Audigier V, Resche-Rigon M. **micemd: multiple imputation by chained equations with multilevel data.** 2019. https://rdrr.io/cran/micemd/man/micemd-package.html (accessed Oct 9, 2019).

18. Rubin DB. **Multiple imputation for nonresponse in surveys**. New York: John Wiley & Sons, 1987.

19. Altman DG, Bland JM. **Diagnostic tests 2: Predictive values**. BMJ 1994; 309(6947): 102.

20. Fleischman RJ, Mann NC, Dai M, et al. **Validating the use of ICD-9 code mapping to generate Injury Severity Scores.** J Trauma Nurs 2017; 24(1): 4-14.

21. Newgard CD, Hedges JR, Diggs B, Mullins RJ. **Establishing the need for trauma center care: anatomic injury or resource use?** Prehosp Emerg Care 2008; 12(4): 451-8.

22. Lerner EB, Drendel AL, Cushman JT, et al. **Ability of the physiologic criteria of the field triage guidelines to identify children who need the resources of a trauma center.** Prehosp Emerg Care 2017; 21(2): 180-4.

23. Newgard CD, Zive D, Holmes JF, et al. **A multisite assessment of the American College of Surgeons Committee on Trauma field triage decision scheme for identifying seriously injured children and adults.** J Am Coll Surg 2011; 213(6): 709-21.

24. Lerner EB, Studnek JR, Fumo N, et al. **Multicenter analysis of transport destinations for pediatric prehospital patients.** Acad Emerg Med 2019; 26(5): 510-6.

25. van Rein EAJ, van der Sluijs R, Voskens FJ, et al. **Development and validation of a prediction model for prehospital triage of trauma patients.** JAMA Surg 2019.

26. Fleming S, Thompson M, Stevens R, et al. **Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies.** Lancet 2011; 377(9770): 1011-8.

27. Paszke A, Gross S, Chintala S, et al. **Automatic differentation in PyTorch.** NIPS Autodiff Workshop 2017.

28. Kingma D, Ba J. **Adam: A method for stochastic optimization.** Proceedings of the 3rd International Conference on Learning Representations 2014.

29. Borg A, Sariyar M. **RecordLinkage: record linkage in R.** 2019.

30. R Core Team. **R: A language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing; 2019.

APPENDIX

SelectAssist

Patient selection in pre-hospital triage is often based on chief complaints registered by Emergency Medical Service (EMS) professionals in electronic health records (EHR). The P2-T2 study, however, first included all patients transported by EMSs from the scene of injury to a participating emergency department independent of chief complaint in order to prevent selection bias. Chief complaints were missing in 29.1% and many patients in need of trauma-related evaluation at the emergency department had a different chief complaint. For example, 19.1% of the patients with a neurologic chief complaint in a sample of patients were considered injured after manual review (e.g., traumatic brain injuries).

This selection strategy resulted in a large dataset including 631,475 patients that needed to be reviewed. We developed a tool (SelectAssist) to aid patient selection for that reason. First, we defined the study domain as patients in need of trauma-related evaluation at the emergency department, irrespective of additional complaints. Two researchers reviewed a combined total of 19,525 EHRs to manually label the need of trauma-related evaluation at the emergency department (further referred to as the label) for each individual patient.

Data of individual records were split into text fields and additional variables. All text fields were merged into one variable. The text variables were anonymized, converted to lowercase, and pre-padded to construct equally sized variables. All manually labelled EHRs were subsequently divided into a training (n=15,814), a validation (n=1758), and a test set (n=1953). A recurrent neural network was developed with the PyTorch and Torchtext software packages to predict the labels.[27] The aforementioned text variables were used as an input and the model's output was the probability of a positive label. The network architecture consisted of an embedding layer, two blocks of long-short term memory (LSTM) layers, and a final fully connected layer. The training phase was performed in 80 epochs, using the Adam optimizer, and a learning rate scheduler.[28]

The best model had an overall accuracy of 98.0% in the training set, 97.3% in the validation set, and 96.9% in the test set. Predictions were generated using this newly developed model for all included EHRs. The resulting predictions were used in conjunction with other non-text variables of the manually labelled EHRs to create the final prediction model (SelectAssist). The non-text variables consisted of, amongst others: age, gender, administered medicine, and chief complaint. Finally, a logistic regression model was developed with the predictions from the recurrent neural network and the

non-text variables as input. Basis splines with three knots were used to fit variables that were considered to have a non-linear relationship with respect to the label.

The overall accuracy of SelectAssist in the hold-out test set was 99% (95%-CI, 98 – 99) with a sensitivity of 97% (95%-CI, 95.0 – 98) and a specificity of 99% (95%-CI, 98 – 100). The c-statistic in the hold-out test set was: 0.99 (95%-CI, 0.99 – 1.0). These results demonstrate that SelectAssist is a very accurate tool to select patients in need of trauma-related emergency department evaluation. The threshold to classify a patient as being injured was lowered to 0.39 and all resulting patients were subsequently manually reviewed for this study. The predicted label was manually switched for 145 of the initially selected patients.

LinkAssist

Pre-hospital EHRs needed to be linked with in-hospital outcomes (e.g., Injury Severity Scores) to analyze triage accuracy based on either decision rules or initial transportation destination. General record linkage is implemented in a couple of software packages in the R statistical programming language.[29,30] However, these packages were either not accurate enough or were computationally infeasible because of the magnitude of the current dataset. To enable accurate, anonymous, and computationally feasible record linkage we developed LinkAssist, specifically for this dataset.

Records were either linked deterministically or probabilistically. Deterministic linkage was performed when a unique patient identifier was available in both pre-hospital and in-hospital data. A total of 28,634 records were deterministically linked. These

Table Definition of Major Surgical Intervention

Major surgical interventions
Damage control thoracotomy
Damage control laparotomy
Extraperitoneal pelvic packing
Revascularization of extremities
Craniotomy
Coniotomy/cricothyrotomy
Damage control orthopedics

One or more of these interventions results in a positive reference standard.

deterministically linked records were then split into a training (n=25,771) and a test set (n=2863) in order to develop a prediction model for probabilistic record linkage. Similar variables were extracted from pre-hospital EHRs and in-hospital data. Included predictors were, amongst others: age, sex, one-way hashes of the day, month, and year of birth, in conjunction with the hashed day and month of injury, the year in which the injury occurred, and the individual components of the Glasgow Coma Scale. Thirty records transported to the same hospital in the same year were randomly sampled from the in-hospital data for every pre-hospital record in the training set. The test set included all available records transported to the same hospital in the same year for each pre-hospital record (i.e., no sampling was performed). A gradient boosting machine was constructed to predict potential matching records in the augmented training set consisting of the pre-hospital records in the training set, the sampled in-hospital data, and the actual matching in-hospital data. Ten-fold cross-validation was used to select appropriate hyperparameters and to prevent model optimism. We assumed that model's predictive accuracy would generalize to the probabilistic dataset as the three predictors (absolute difference in age, equal day of injury, and equal day of birthday) that yielded over 98% of the information gain were similarly distributed in the deterministically linked data and had no missing values..

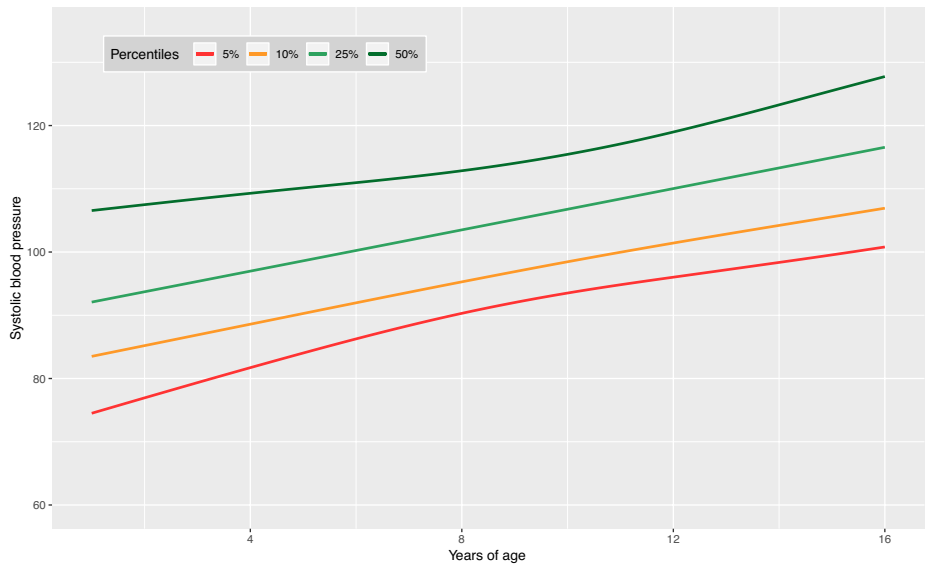


Figure Percentile Chart of Systolic Blood Pressure for Children <16 Years of Age. *Percentiles were estimated using quantile regression.

The resulting model proved to be highly accurate in the hold-out test set (100%) with a sensitivity of 100% (95%-CI, 99 – 100) and specificity of 100% (95%-CI, 100 – 100). The c-statistic in the test set was 1.0 (95%-CI, 1.0 – 1.0). This model (LinkAssist) was subsequently used to predict potential matches for all non-deterministically linked records in the cohort. All hospitalized patients with a positive primary or secondary reference standard in the Dutch National Trauma Registry were manually reviewed to prevent information bias.

CHAPTER IV

Accuracy of pre-hospital triage in selecting severely injured trauma patients

JAMA Surgery

Frank J. Voskens; Eveline A.J. van Rein; Rogier van der Sluijs;
Roderick M. Houwert; Robert A. Lichtveld; Egbert J. Verleisdonk; Michiel Segers;
Ger van Olden; Marcel Dijkgraaf; Luke P.H. Leenen; Mark van Heijl

ABSTRACT

Importance

A major component of trauma care is adequate pre-hospital triage. To optimize the pre-hospital triage system, it is essential to gain insight in the quality of pre-hospital triage of the entire trauma system.

Objective

To prospectively evaluate the quality of the field triage system to identify severely injured adult trauma patients.

Design, Setting, and Participants

Pre-hospital and in-hospital data of all adult trauma patients during 2012 to 2014 transported with the highest priority by Emergency Medical Services professionals to 10 hospitals in Central Netherlands were prospectively collected. Pre-hospital data collected by the Emergency Medical Services professionals were matched to hospital data collected in the trauma registry. An Injury Severity Score of 16 or more was used to determine severe injury.

Main Outcomes and Measures

The quality and diagnostic accuracy of the field triage protocol and compliance of Emergency Medical Services professionals to the protocol.

Results

A total of 4950 trauma patients were evaluated of which 436 (8.8%) patients were severely injured. The undertriage rate based on actual destination facility was 21.6% (95%-CI, 18.0 – 25.7) with an overtriage rate of 30.6% (95%-CI, 29.3 – 32.0). Analysis of the protocol itself, regardless of destination facility, resulted in an undertriage of 63.8% (95%-CI, 59.2 – 68.1) and overtriage of 7.4% (95%-CI, 6.7 – 8.2). The compliance to the field triage trauma protocol was 73% for patients with a level-I indication.

Conclusions and Relevance

More than 20% of the patients with severe injuries were not transported to a level-I trauma center. These patients are at risk for preventable morbidity and mortality. This finding indicates the need for improvement of the pre-hospital triage protocol.

INTRODUCTION

Adequate pre-hospital trauma triage of injured patients is imperative for optimal trauma care. In an inclusive trauma system, it is essential to transport patients with severe injuries to a level-I trauma center and patients without severe injuries to lower-level hospitals.[1,2] Previous studies have clearly shown lower mortality rates in patients with severe injuries treated at a level-I trauma center compared with patients treated at a lower-level hospital.[1-6]

Management of care of the injured trauma patient on scene remains challenging, and situations can be chaotic. After a rapid trauma assessment of clinical and physiologic parameters, Emergency Medical Services (EMS) professionals must identify patients at risk for severe injury and select the proper destination. Pre-hospital triage protocols are used to help define the patient destination. However, triage of patients without evident abnormality and instability at presentation remains challenging given the limited facilities on scene.

In the Netherlands, allocation of trauma patients to the appropriate level of trauma care is guided by the Dutch field triage protocol (version 7.1, National Protocol of Ambulance Services), for EMS professionals (Figure 1).[7] This protocol is based on the Field Triage Decision Scheme established by the American College of Surgeons Committee on Trauma (ACSCOT).[8,9]

Quality of pre-hospital triage can be determined by rates of undertriage and overtriage. Undertriage is defined as the proportion of patients with severe injuries not transported to a level-I trauma center. Overtriage is defined as the proportion of patients without severe injuries transported to a level-I trauma center. Undertriage results in higher mortality and delay of adequate care, whereas overtriage limits the available level-I resources for patients without severe injuries.[2, 8] To optimize the pre-hospital triage system, it is essential to gain insight in the quality of pre-hospital triage of the entire trauma system or region. The benchmark level in the ACSCOT guidelines is a maximum undertriage rate of 5%, allowing for an overtriage rate of up to 50%.[8] In a Dutch population consisting of high-energy trauma patients only, the undertriage rate was 11%.[10] The quality of triage in the complete trauma population is however unknown.

This present study aims to evaluate the quality of the Dutch field triage protocol for identifying severely injured trauma patients in a population consisting of adult trauma patients transported by EMS professionals with the highest priority in the region Central Netherlands.

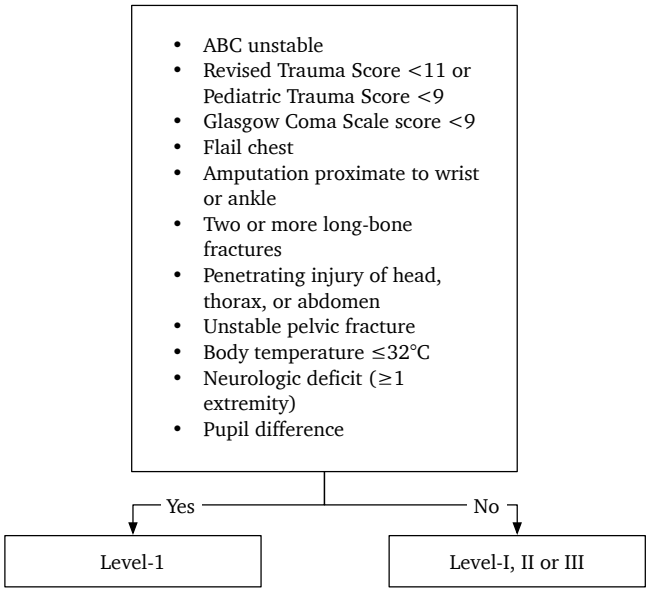


Figure 1 Abstraction of The National Protocol of Ambulance Services (Triage Choice of Hospital, version 7.1)

METHODS

Study Design and Setting

The present study was performed in Central Netherlands region using prospectively collected pre-hospital and in-hospital data of all adult trauma patients transported with the highest priority by the Regional Ambulance Service Utrecht to one of 10 hospitals in Central Netherlands between January 2012 and July 2014. The region Central Netherlands consists of nine level-II or III hospitals and one level-I trauma center in a 2418 km^2 region with a population of 1.2 million people. The University Medical Center Utrecht is designated as a level-I trauma center, offering trauma care at the highest level for severely injured patients. The nine surrounding level-II or III hospitals are designated to treat patients without severe injuries. This regional trauma network is based on an inclusive and integrated trauma system.[8] The ambulance care system is nurse based. Ambulance nurses are licensed to administer medical treatment at advanced life support level and ambulance drivers are qualified to provide medical assistance to the ambulance nurses. The present study protocol was reviewed and approved by the local Medical Ethical Committee, and patient consent was waived. Analyses began in 2016.

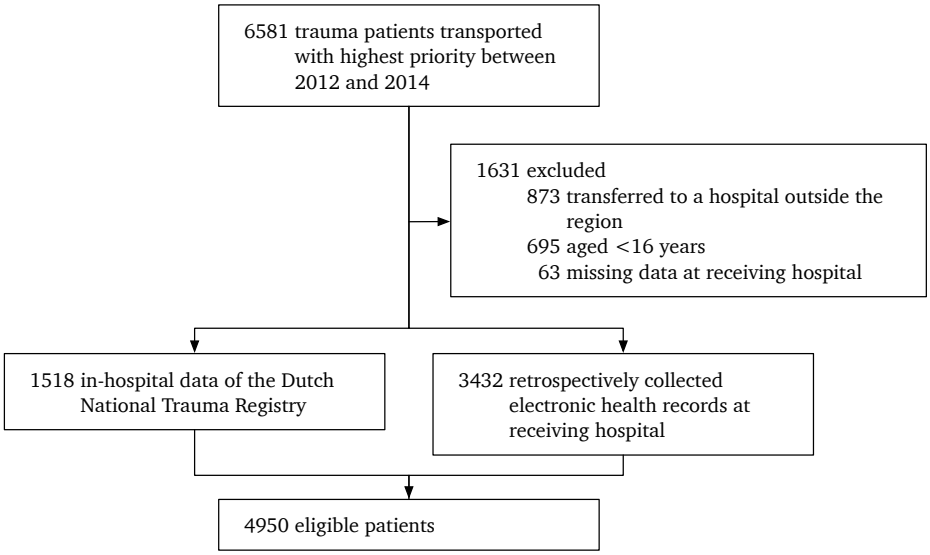


Figure 2 Study Profile

Patients

All trauma patients aged 16 years and older transported by EMS professionals with the highest priority were included in the study. Patients transported to a hospital outside the region Central Netherlands and patients transported by helicopter were excluded. Patients were also excluded if insufficient data were available in the receiving hospital to properly calculate the Injury Severity Score (ISS).

Data Collection

Pre-hospital reports from the EMS professionals were prospectively collected and included patient demographics, description of the trauma mechanism, physical examination data on site, pre-hospital treatment, and receiving hospital. Furthermore, the report included a standardized digital report of specific vital parameters: i.e., Glasgow Coma Scale score, respiratory rate, systolic blood pressure, heart rate, pupil deficit, and the Revised Trauma Score.

The Dutch National Trauma Registry registers in-hospital data regarding injuries and complications for all trauma patients admitted to a hospital. For patients who were discharged from the emergency department, data were extracted from the electronic patient documentation. Injuries were encoded according to the Abbreviated Injury Scale 90 Update 98 (AIS 98).[11] ISS scores were calculated and used to assess overall injury severity.

Outcome

Severe injury was defined as an Injury Severity Score (ISS) of 16 or more. The primary outcome of this study was the quality of the field triage system in terms of undertriage and overtriage. Undertriage was defined as the proportion of severely injured patients (ISS ≥16) erroneously transported to level-II or III hospitals. Overtriage was defined as the proportion of patients with an ISS of less than 16 transported to a level-I trauma center.[8,12,13]

The diagnostic accuracy of the Dutch field triage protocol was calculated for identifying patients with or without severe injuries, regardless of actual destination facility. For this purpose, the level-I triage criteria were retrospectively applied to the dataset. For this part of the analysis, undertriage was defined as the proportion of patients with severe injuries not identified by the pre-hospital trauma triage protocol, divided by the total number of severely injured patients. Overtriage was defined as the proportion of patients without severe injuries identified as severely injured patients using the pre-hospital trauma triage protocol. Pre-hospital level-I criteria were penetrating injury (head, thorax and/or abdomen), two or more fractures of long bones (humerus and/or femur), amputation proximate to wrist or ankle, neurologic failure in one or more extremity, unstable pelvic fracture, pupil difference, flail chest, Glasgow Coma Scale score more than 9, deteriorating Glasgow Coma Scale score, Revised Trauma Score less than 11, vitally compromised in airway and breathing or circulation, and body temperature of 32°C and less.

Finally, the compliance of EMS professionals for correct transportation of patients with pre-hospital level-I trauma center criteria according to the Dutch field triage protocol was determined.

Statistical Analysis

Data were analyzed using descriptive statistics and results were shown in frequencies, and percentages. Undertriage and overtriage rates were presented with 95%-CI. Multiple imputation was used for missing pre-hospital values and was performed with SPSS IBM statistical software (version 23.0). Missing values were predicted based on all other predictors, as well as the outcome (ISS). All logistic regression analyses were performed on five imputed datasets independently and pooled afterwards for missing pre-hospital values. Multiple imputation for missing pre-hospital values has been previously validated.[14] Multiple imputation was used for: pulse in 6.8%, respiratory rate 6.5%, systolic blood pressure 7.0%, Revised Trauma Score in 8.1%, and Glasgow Coma Scale score in 4.6%.

RESULTS

A total of 6581 trauma patients were transported by EMS professionals with the highest priority in Central Netherlands. Inclusion criteria were met in 4950 patients for the current analysis (Figure 2). Characteristics of the study sample are shown in Table 1. Patients were relatively equally distributed between the hospitals: level-I, 1724 (34.8%) patients; level-II, 1326 (26.8%) patients; and level-III, 1900 (38.4%) patients. Median age was 45 years, 2887 were male (58.3%), and 436 (8.8%) patients had an ISS of 16 or more. Severe injury in one of the body regions (Abbreviated Injury Scale score ≥3) was most frequently diagnosed in the head and extremities.

Table 1 Characteristics of the Study Population

Variable	**Patients, no. (%) (n=4950)**
Male	2887 (58.3%)
Age in years, median (IQR)	45 (27 – 63)
Elderly adults (>65 years)	1085 (21.9%)
Pre-hospital Glasgow Coma Scale score <9	141 (2.8%)
Triage criteria	
Mechanism of injury	1300 (26.3%)
Physiologic criteria	289 (5.8%)
Injury criteria	256 (5.2%)
Assistance of air medical services	119 (2.4%)
Out-of-hospital intubation	49 (1%)
Transfer to	
Level-I trauma center	1724 (34.8%)
Level-II trauma center	1326 (26.8%)
Level-III trauma center	1900 (38.4%)
ISS, median (IQR)	2 (5)
Severely injured (ISS ≥16)	436 (8.8%)
AIS score ≥3, per region	
Head and neck	435 (8.8%)
Face	26 (0.5%)
Thorax	318 (6.4%)
Abdomen	61 (1.2%)
Extremities	496 (10%)
External	12 (0.2%)
In-hospital stay	2047 (41.2%)
Mortality	61 (1.2%)

Abbreviations: AIS, *Abbreviated Injury Scale*; ISS, *Injury Severity Score*; IQR, *Inter Quartile Range*.

Table 2 Quality of Field Triage System Regarding Correct Destination Facility for Patients with and without Severe Injuries

Variable	Patients with severe injuries (ISS ≥16) (n=436)	Patients without severe injuries (ISS <16) (n=4514)
Quality of field triage system regarding correct destination facility for patients with and without severe injuries		
Level-I trauma center, no. (%)	342 (78.4%)	1382 (30.6%)
Level II/III trauma center, no. (%)	94 (21.6%)	3132 (69.4%)
Undertriage, % (95%-CI)[a]	21.6 (18 – 25.7)	NA
Overtriage, % (95%-CI)[b]	NA	30.6 (29.3 – 32.0)
Diagnostic accuracy of the Dutch pre-hospital field triage protocol for identifying patients with and without severe injuries		
Level-I trauma center indication, no. (%)	158 (36.2)	334 (7.4)
No level-I trauma center indication, no. (%)	278 (63.8)	4180 (92.6)
Undertriage, % (95%-CI)[c]	63.8 (59.2 – 68.1)	NA
Overtriage, % (95%-CI)[d]	NA	7.4 (6.7 – 8.2)

Abbreviations: ISS, *Injury Severity Score*; NA, *not applicable*.
[a] Percentage of patients with severe injuries (ISS ≥16) not transported to level-I trauma center.
[b] Percentage of patients without severe injuries (ISS <16) transported to level-I trauma center.
[c] Percentage of patients with severe injuries (ISS ≥16) without positive pre-hospital level-I criteria according to the field triage protocol.
[d] Percentage of patients without severe injuries (ISS <16) with positive pre-hospital level-I criteria according to the field triage protocol.

Of 436 patients with severe injuries, 94 were erroneously transported to level-II or III hospitals, resulting in an undertriage of 21.6% (95%-CI, 18.0 – 25.7). Transportation of 1382 of 4514 patients without severe injuries to the level-I trauma center resulted in an overtriage of 30.6% (95%-CI, 29.3 – 32.0; Table 2).

The diagnostic accuracy of the Dutch field triage protocol is shown in Table 2. The protocol-based undertriage was 63.8% (95%-CI, 59.2 – 68.1) while the protocol based overtriage was 7.4% (95%-CI, 6.7 – 8.2). The compliance of EMS professionals to the field triage protocol was 72.6%. Overall, 30% percent of the patients with a positive injury and/or physiology criteria were not transported to a level-I trauma center.
Table 3 illustrates the undertriage and overtriage rates for different subgroups of the study population regarding correct destination facility. The undertriage rate in elderly patients is high at 38.6% (95%-CI, 30.8 – 47.2). A high energy trauma mechanism resulted in an

Table 3 Quality of Field Triage Systems Regarding Correct Destination Facility for Different Subgroups

Variable	Total patients, no.	Patients with severe injuries, no. (%)	Undertriage (95%-CI)	Overtriage (95%-CI)
Men	2887	295 (10.2)	19.7 (15.5 – 24.6)	32.8 (31 – 34.6)
Women	2063	141 (6.8)	25.5 (19.1 – 33.3)	27.7 (25.8 – 29.8)
Adults (≤65 years)	3865	304 (7.9)	14.1 (10.7 – 18.5)	33.2 (31.6 – 34.7)
Elderly adults (>65 years)	1085	132 (12.2)	38.6 (30.8 – 47.2)	21.1 (18.6 – 23.8)
Mechanism criteria	1301	186 (14.3)	9.1 (5.8 – 14.2)	55.3 (41.8 – 47.6)
Injury criteria	256	81 (31.6)	0 (0 – 4.5)	33.1 (26.6 – 40.4)
Physiologic criteria	289	116 (40.1)	2.6 (0.9 – 7.3)	55.5 (48.1 – 62.7)
Head injury	2143	304 (14.2)	22 (17.7 – 27.1)	32.7 (30.6 – 34.9)

undertriage rate of 9.1% (95%-CI, 5.8 – 14.2). The group of patients with a positive injury and/or physiologic criteria showed low undertriage rates (0% and 2.6%, respectively).

DISCUSSION

This study presents a quality assessment of pre-hospital triage in identifying severely injured trauma patients using prospectively collected data. Pre-hospital data were collected from EMS professionals and included every type of trauma patient transported with the highest priority, whether admitted or discharged from the emergency department in all types of hospitals.

The quality of the Dutch field triage system remains relatively low. The overall rate of undertriage of the pre-hospital trauma triage system was 22% and is significantly higher than the benchmark level of 5%, as set by the ACSCOT.[15] This implies that a significant group of trauma patients with severe injuries does not receive the appropriate level-I trauma care. These patients are therefore at risk for increased morbidity and mortality.[6,8,16] A variety of causes can be identified for undertriage. Closer examination of the elderly patients (age ≥65 years) in the present study showed a high undertriage rate of 39%. The undertriage rate among the elderly patients was 25% higher compared with younger adults. These findings are in accordance with previous studies showing increased undertriage rates in elderly patients.[17,18] Elderly patients tend to have more cognitive and physical impairments with pre-existing co-morbidity and therefore low energy trauma mechanisms may result in serious injuries.[18] Undertriage of elderly patients remains a

substantial problem. Modifications to the adult criteria of the ACSCOT triage protocol have been made to accentuate these physiologic and anatomic differences of the elderly population.[19] However, the effect of these modifications has not yet been evaluated.

A considerable proportion of trauma patients in our study population experienced traumatic brain injury. Patients with traumatic brain injury are at risk of undertriage because the identification of significant traumatic brain injury can be demanding in the pre-hospital setting. Previous studies showed a high risk of undertriage in patients with isolated head injuries.[20,21] In our study population, more than 75% of the undertriaged patients received a diagnosis of a cranial Abbreviated Injury score of at least 3. These patients need access for direct neurosurgical care.

Overtriage is also an important outcome parameter to monitor because high overtriage rates lead to reduced system efficiency, unnecessary burden to the level-I trauma center, and lower cost-effectiveness.[22,23] Our study showed an acceptable overtriage rate of 31% (95%-CI, 29.3 – 32.0). Overtriage rates of up to 50% are acceptable for pre-hospital triage to keep undertriage rates to a minimum.[8,15] Therefore, an overtriage rate of 31% should provide room for improvement of undertriage rates.

Our findings support the results of recent studies and confirm that undertriage rates remain high. Although past studies were retrospective, one exception is a 2016 prospective study investigating the ACSCOT triage protocol including 17,633 trauma patients, of which 3% were seriously injured (ISS ≥16).[17] A large group of patients was excluded because of study sampling design, and Newgard and colleagues reported a significant amount of missing hospital data. The authors reported an undertriage rate of 36.4% and an overtriage rate of 28.7% based on the initial receiving hospital. After accounting for inter-hospital transfers, the undertriage rate was 22%. This higher undertriage rate compared to our results cannot be explained by the difference in protocols because the Dutch triage protocol has a higher threshold for transportation to a level-I trauma center. However, there are significant regional differences. Hospitals in the region Central Netherlands are clustered in a relatively close proximity. The level-I trauma center is always within a 15-minute drive for an ambulance, whereas this could be more than 60 minutes in some of the regions studied by Newgard et al. The significant lowering of the undertriage rate after accounting for inter-hospital transfers in the study of Newgard and colleagues could suggest at least some role of hospital proximity. Previous research shows higher mortality rates in trauma patients after inter-hospital transfers. This emphasizes the need to correctly identify and transport patients with severe injuries directly to a level-I trauma center.[16]

A previous evaluation of the Dutch field triage by our study group revealed undertriage and overtriage rates of 10.9% and 39.5%, respectively.[10] However, this study exclusively included high-energy trauma patients, which could very well explain the difference in Dutch field triage protocol rates compared with the present study. Although high-energy trauma is not a strict level-I criterion in the current field triage protocol, it can be hypothesized that patients who have an obvious high energy trauma are more prone to be transported to a level-I trauma center owing to EMS professionals' judgment. Other studies have demonstrated that the use of a mechanism of injury criterion could lower undertriage and suggest that specific high-energy trauma criteria should be included in the level-I criteria.[24] Evaluation of the subgroup of patients after a high-energy trauma in the present study also revealed a lower undertriage rate of 9%, supporting the suggestion to include mechanism of injury as level-I criterion.

The exceptionally high undertriage rate calculated for the pre-hospital protocol itself reflects the shortcomings of the currently used protocol. It truly fails to support EMS professionals to correctly identify severely injured patients in need of level-I trauma care. Fortunately, owing to EMS professionals' judgment, a large group of the severely injured trauma patients was still transported to the right trauma center and received appropriate care. The discrepancy in undertriage rates between the protocol itself (64%) and actual undertriage based on destination facility (22%) is probably best explained by the correct assessment of the EMS professionals based on experience regardless of the triage protocol. Previous studies also showed improved triage rates after including EMS professionals' judgment as a triage criterion.[25,26] The preference of the patient and existing transport patterns could also influence the decision for destination facility; however, the impact of these factors could not be assessed.

Strengths and Limitations

A strength of this study is the prospective pre-hospital data collection and study design that included all trauma patients transported to all types of hospitals in a specific region. Furthermore, the triage protocol currently investigated is based on the ACSCOT triage protocol, which is adapted as a standard in many organizations worldwide. Therefore, the findings of this study may be of international importance.

This study has several limitations. The exclusion of patients transported to hospitals outside the study region could result in sampling bias, the extent of which is unknown. As mentioned, the Dutch National Trauma Registry does not register patients who are not admitted to the hospital. Although data of admitted trauma patients were prospectively collected by a dedicated group of trained data managers, our research group collected the hospital data of patients discharged from the emergency department retrospectively.

These retrospectively collected data are not expected to affect the study results because this group consists of patients with minor injuries.

Conclusion

The present study shows that more than 20% of the patients with severe injuries were not transported to a level-I trauma center. These patients are at risk for preventable morbidity and mortality. It also showed that the accuracy of the Dutch field triage protocol in selecting patients with severe injuries is low and therefore of insufficient help to EMS professionals. Our findings indicate the need for improvement of the pre-hospital triage protocol.

REFERENCES

1. Twijnstra MJ, Moons KG, Simmermacher RK, Leenen LP. **Regional trauma system reduces mortality and changes admission rates: a before and after study.** Ann Surg 2010; 251(2):339-43.

2. MacKenzie EJ, Rivara FP, Jurkovich GJ, Nathens AB, Frey KP, Egleston BL, Salkever DS, Scharfstein DO. **A national evaluation of the effect of trauma-center care on mortality.** NEJM 2006; 354(4):366-78.

3. Celso B, Tepas J, Langland-Orban B, Pracht E, Papa L, Lottenberg L, Flint L. **A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems.** J Trauma 2006; 60(2):371-8.

4. Newgard CD, Zive D, Holmes JF, et al. **A multisite assessment of the American College of Surgeons Committee on Trauma field triage decision scheme for identifying seriously injured children and adults.** Journal of the American College of Surgeons. 2011; 213(6):709-21.

5. Sampalis JS, Denis R, Lavoie A, Frechette P, Boukas S, Nikolis A, Benoit D, Fleiszer D, Brown R, Churchill-Smith M, et al. **Trauma care regionalization: a process-outcome evaluation.** J Trauma 1999; 46(4):565-79.

6. Sampalis JS, Denis R, Frechette P, Brown R, Fleiszer D, Mulder D. **Direct transport to tertiary trauma centers versus transfer from lower-level facilities: impact on mortality and morbidity among patients with major trauma.** J Trauma 1997; 43(2):288-95.

7. Ambulancezorg Nederland. **Landelijk Protocol Ambulancezorg (versie 7.1)**. Zwolle, 2016.

8. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources.

9. Mackersie RC. **History of trauma field triage development and the American College of Surgeons criteria.** Prehosp Emerg Care 2006; 10(3):287-94.

10. van Laarhoven JJ, Lansink KW, van Heijl M, Lichtveld RA, Leenen LP. **Accuracy of the field triage protocol in selecting severely injured patients after high energy trauma.** Injury 2014; 45(5):869-73.

11. Gennarelli T, Wodzin E. **The Abbreviated Injury Scale 2005: update 2008.** Association for the Advancement of Automotive Medicine; Des Plaines, IL: 2008.

12. Champion HR, Copes WS, Sacco WJ, et al. **The Major Trauma Outcome Study: establishing national norms for trauma care.** J Trauma 1990; 30(11):1356-65.

13. Baker SP ONB, Haddon WJr, Long WB. **The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care.** J Trauma 1974; 14:187-96.

14. Newgard CD, Haukoos JS. **Advanced statistics: missing data in clinical research--part 2: multiple imputation.** Acad Emerg Med 2007; 14(7):669-78.

15. American College of Surgeons Committee on Trauma. **Hospital and pre-hospital resources for optimal care of the injured patient.** Bull Am Coll Surg 1986; 71(10):4-23.

16. Mans S, Reinders Folmer E, de Jongh MA, Lansink KW. **Direct transport versus inter**

hospital transfer of severely injured trauma patients. Injury 2016; 47(1):26-31.

17. Newgard CD, Fu R, Zive D, et al. **Prospective validation of the national field triage guidelines for identifying seriously injured persons.** JACS 2016; 222(2):146-58.e2.

18. Newgard CD, Richardson D, Holmes JF, et al. **Physiologic field triage criteria for identifying seriously injured older adults.** Prehosp Emerg Care 2014; 18(4):461-70.

19. Sasser SM, Hunt RC, Faul M, et al. **Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011.** MMWR 2012; 61(Rr-1):1-20.

20. Nakahara S, Matsuoka T, Ueno M, et al. **Predictive factors for undertriage among severe blunt trauma patients: what enables them to slip through an established trauma triage protocol?** J Trauma 2010; 68(5):1044-51.

21. Xiang H, Wheeler KK, Groner JI, Shi J, Haley KJ**. Undertriage of major trauma patients in the US emergency departments.** Am J Emerg Med 2014; 32(9):997-1004.

22. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

23. MacKenzie EJ, Weir S, Rivara FP, et al. **The value of trauma center care.** J Trauma 2010; 69(1):1-10.

24. Brown JB, Stassen NA, Bankey PE, Sangosanya AT, Cheng JD, Gestring ML. **Mechanism of injury and special consideration criteria still matter: an evaluation of the National Trauma Triage Protocol.** J Trauma 2011; 70(1):38-44; discussion -5.

25. Lavoie A, Emond M, Moore L, Camden S, Liberman M. **Evaluation of the Pre-hospital Index, presence of high-velocity impact and judgment of emergency medical technicians as criteria for trauma triage.** Cjem 2010; 12(2):111-8.

26. Newgard CD KM, Nelson M, et al. **Deciphering the use and predictive value of "EMS Provider Judgment" in out-of-hospital trauma triage: a multi-site, mixed methods assessment.** J Trauma 2012; 72(5):1239-48.

27. Pape HC, Lefering R, Butcher N, et al. **The definition of polytrauma revisited: an international consensus process and proposal of the new 'Berlin definition'.** J Trauma 2014; 77(5):780-6.

28. Wong TH, Lumsdaine W, Hardy BM, Lee K, Balogh ZJ. **The impact of specialist trauma service on major trauma mortality.** J Trauma 2013; 74(3):780-4.

29. Baker SP ONB, Haddon W Jr, et al. **The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care.** J Trauma 1974; 14:187–96.

PART II

The trauma triage continuum of care cohort

CHAPTER V

Cohort Profile: the Trauma Continuum of Care Cohort (TRACCC)

Rogier van der Sluijs; Job F. Waalwijk; Robin D. Lokerman;
Martijn Poeze; Luke P.H. Leenen; Mark van Heijl;
for the Pre-hospital Trauma Triage Research Collaborative

ABSTRACT

Cohort Purpose

The Trauma Continuum of Care Cohort (TRACCC) was designed to study the effects of pre-hospital care and decision-making on in-hospital patient outcomes. It covers data on the decision-making and transportation destinations of Emergency Medical Services (EMS) within seven inclusive trauma regions. The cohort was designed to evaluate the accuracy of field triage in both adult and pediatric populations, and to develop and validate prediction models to aid EMS professionals in selecting patients in need of specialized trauma care in the pre-hospital setting.

Cohort Basics

Baseline data of 165,404 trauma patients transported by ground ambulances of eight Emergency Medical Services to 70 hospitals within seven inclusive trauma regions were collected between 2015 and 2017.

Design and Measures

TRACCC is a prospective cohort study constructed from individual participant data from multiple Emergency Medical Services linked to in-hospital patient outcomes. It contains demographic characteristics, vital signs, mechanisms of injury, and injury types, and a diverse set of patient outcomes including injury severity metrics, resource use, and mortality. Additional data included in TRACCC allow researchers to conduct geospatial analyses.

Unique Features

Unique features of TRACCC are the robust and sensitive methodology used to identify patients, the highly accurate record linkage strategy, and enormous scale. Participating EMSs serve urban, suburban, and rural regions with heterogeneous patient characteristics, varying geography, and different hospital density, enabling researchers to evaluate the generalizability of their findings.

Reasons To Be Cautious

Missing data is empirical in time-critical settings such as field triage. Consequently, vital signs are missing up to 36%. Although these values are illustrative of the pre-hospital setting, it often requires complex (multi-level) multiple imputation strategies before data can be analyzed. In addition, there is a low proportion of severe injuries because of relatively broad inclusion criteria.

Collaboration and Data Access

The scientific committee of TRACCC welcomes proposals for new research, collaboration in data analysis, data interpretation, and publication. The scientific committee can be contacted through the web form available at https://www.traumatriageapp.com/en/.

WHY WAS THE COHORT SET UP?

The major goal of inclusive trauma systems is to facilitate optimal care along a continuum that ranges from the pre-hospital setting to rehabilitation.[1] Centralization and integration of resources within such systems is fundamental to enable optimal care. The greater the amount of centralization, the greater the importance of adequate field triage. Field triage is the diagnostic strategy that aims to get the right patient to the right hospital in the right time.[1] First, Emergency Medical Services (EMS) professionals need to assess the patient's resource-need at the scene of injury. Second, facilities with matching resources must be identified, given constraints such as trauma center proximity, capacity, and patient acuity. Finally, the optimal transportation destination has to be determined in light of steps one and two.

Mistriage in trauma systems could be disastrous. Transporting severely injured patients to lower-level trauma centers (i.e., undertriage) is a medical problem associated with increased mortality rates, whereas overtriage – transporting mildly injured patients to higher-level trauma centers – is associated with disproportionate costs and exhaustive use of scarce resources.[2-4]

A recent series of systematic reviews uncovered mistriage to be a worldwide problem in both the adult and pediatric population.[5-7] No inclusive trauma system nor triage tool was identified that could attain an undertriage rate of <5% in combination with acceptable overtriage rates (as is recommended by the American College of Surgeons Committee on Trauma).[1]

The Dutch Trauma Continuum of Care Cohort (TRACCC) set out to study pre-hospital trauma triage, including, but not limited to, the accuracy of triage protocols, the accuracy of triage based on the initial transportation destination, and compliance of EMS professionals to triage tools. TRACCC is composed of different data sources, from multiple regions, including pre-hospital electronic health records, resource use, and patient outcomes. The volume, variety, and velocity of the cohort's data provide novel opportunities for triage research in general and predictive modelling of effects that transcend the pre-hospital setting in particular.

The research questions that TRACCC aims to address are: (i) which combination of – and interaction between – risk factors yield prediction models with optimal model performance for predicting injury severity and in-hospital resource use of trauma patients during field triage?; (ii) what is the diagnostic accuracy of pre-hospital triage of trauma patients transported from the scene of injury to a trauma-receiving emergency department?; (iii) what is the diagnostic accuracy of – and compliance to – triage tools for identifying patients in need of specialized care during field triage?

The Gradient Boosted Trauma Triage (GOAT) study addresses the first research question. This study employs an internal-external cross-validation strategy to develop prediction models with minimal model optimism in order to maximize generalizability. Model performance will be assessed through individual participant data meta-analyses. Finally, the newly developed model will be compared to a previously developed model that is evaluated by EMS professionals in daily practice during field triage in the Netherlands.[8] We partially explored question two and three in an adult population in a one inclusive trauma region in the Netherlands, leaving it unclear to which extent the findings will generalize in other inclusive trauma regions.[9] The inclusion of diverse populations in TRACCC finally provides the opportunity to continuously monitor, research, and optimize field triage on a national scale.

WHO IS IN THE COHORT?

Participating Regions and Hospitals

TRACCC was designed to prospectively collect pre-hospital data, in-hospital resource use, and patient outcomes from January 1, 2015 until December 31, 2017. Patients transported by eight EMSs to 70 hospitals in seven out of 11 inclusive trauma regions in the Netherlands were eligible for inclusion. Participating EMSs serve a mix of rural, suburban, and urban areas (Amsterdam-Amstelland, Zaanstreek-Waterland, Rotterdam-Rijnmond, Zuid-Holland Zuid, Gelderland Zuid, Brabant Midden-West, Brabant Noord, and Utrecht; Figure 1).

Included Patients

Approximately 1.5 million patients were identified through anonymized EMS patient records. Non-urgently transported patients, inter-hospital transfers, patients not transported to a hospital, and duplicated patient records were excluded first, after which 631,475 patients remained eligible.

Trauma patients were automatically identified by an ensemble machine-learning model that was specifically created for this purpose and was named *SelectAssist*. First, a Recurrent

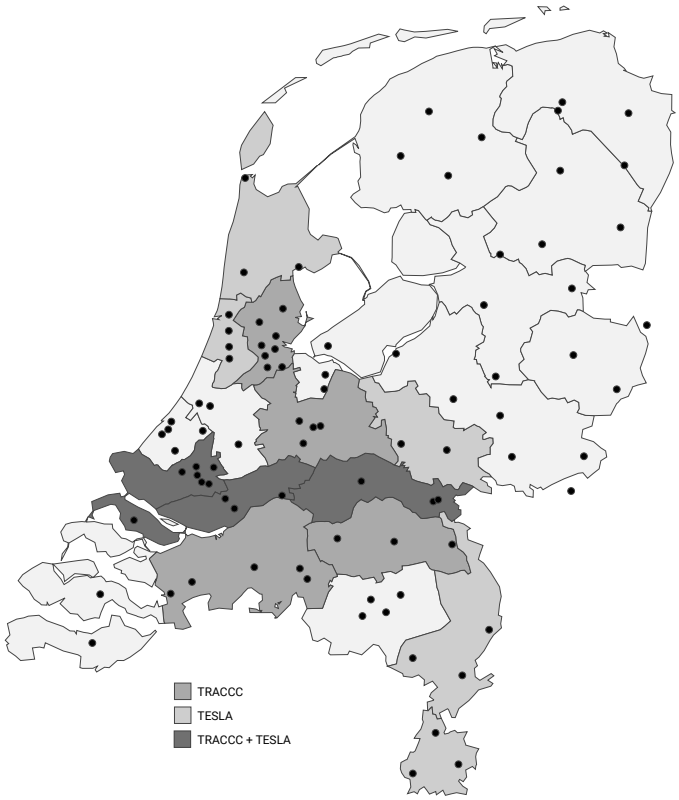


Figure 1 Participating Regions in the Trauma Continuum of Care Cohort (TRACCC)

Neural Network was trained on 15,814 hand-labelled patient records to identify trauma patients based on diagnostic text fields. A learning rate scheduler that decreased on stagnating validation-accuracy decrease was used in conjunction with the Adam optimizer to develop this Neural Network in 50 epochs.[10] Model parameters from each epoch were validated on unseen data from 1758 patients after which the model with the highest validation accuracy was selected. Finally, we evaluated *SelectAssist* on a hold-out test dataset of 1953 patients on which it demonstrated excellent predictive ability.

The predicted probabilities from the Neural Network were used as a predictor in a Logistic Regression model, in conjunction with other predictors such as vital signs, mechanism of injury, injury types, and administered medication. Predictors that were expected to be non-linearly related to the outcome were modelled using basis splines with three knots. The resulting model was developed based on data from 15,814 hand-labelled patient records. Validation on the same hold-out test dataset (n=1953) as

previously described indicated excellent predictive ability with a c-statistic of 0.99 (95%-CI, 0.99 – 1.0) and an overall accuracy of 99% (95%-CI, 98 – 99).

Finally, 165,404 patients were included after computerized exclusion of non-trauma patients and duplicated patient records. The full patient selection strategy is depicted in Figure 2.

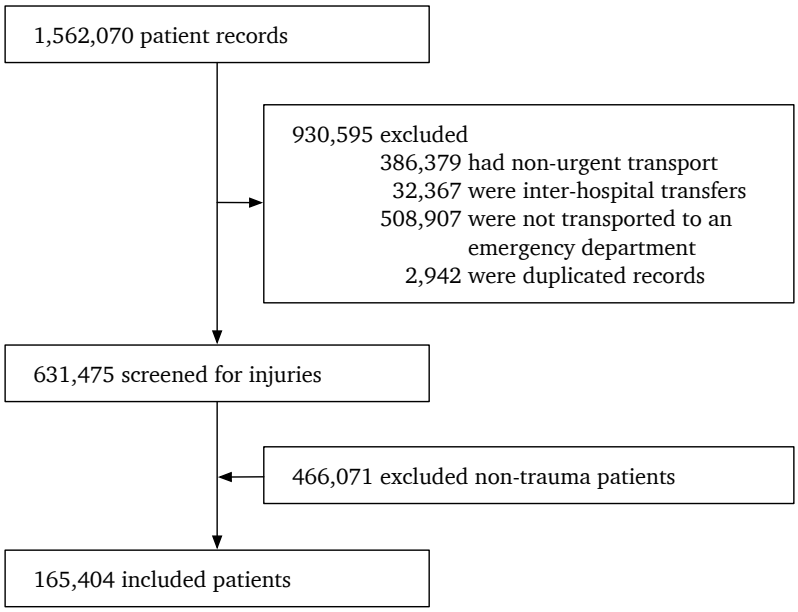


Figure 2 Study Profile

Recruitment from 2017-2020

The design and cohort construction methodology of TRACCC is re-used after 2017 by the Trauma Triage using Supervised Learning Algorithms (TESLA) trial.[11] The TESLA trial is a stepped-wedge, multi-site, multi-center, cluster-randomized trial in which eight EMSs gradually implement a smartphone application that incorporates a recently developed prediction model to identify severely injured patients at the scene of injury.[12] This unidirectional crossover trial aims to explore the impact of this new triage algorithm on the diagnostic accuracy of field triage. The primary endpoint is undertriage, while secondary endpoints include overtriage, hospital resource use, and a cost-utility analysis. Participating EMSs, patient selection, and data collection largely overlaps with the design of TRACCC.

HOW OFTEN HAVE THEY BEEN FOLLOWED UP?

Demographics, resource use, patient outcomes, and other data was collected for all included hospitalized patients by the Dutch National Trauma Registry (in Dutch, *Landelijke Traumaregistratie* [LTR]). The LTR is a nationwide trauma registry that collects in-hospital patient data for every hospitalized trauma patient in the Netherlands. The LTR collects various variables similar to the Utstein template for uniform reporting of data following major trauma.[13]

There was no follow-up of patients discharged from the emergency department. These patients were assumed not to have used any higher-level trauma center resource (e.g., damage control surgery), nor to be severely injured. We verified that no severely injured patients were discharged home from the emergency department in a previous cohort.[9]

The onset of the TESLA trial introduced active follow-up of patients in at least one region. Hospitalized patients consenting to follow-up are sent questionnaires at three, six, 12, and 24 months after the date of injury.

WHAT HAS BEEN MEASURED?

Pre-Hospital Data

Pre-hospital data was collected in a standardized manner from anonymized EMS patient records. Pre-hospital electronic health records used by participating EMSs all included the mandatory variables as specified by the template of the Basic Set of Ambulance Care (in Dutch, *Basisset Ambulancezorg*). This template includes demographics, physiologic characteristics, mechanism of injury, suspected injuries, on-scene treatment, administered medication, and more. Repeated measurement of vital signs such as systolic blood pressure, pulse, and respiratory rate were summarized to their respective first, last, minimal, maximal, and mean measurement values. Information regarding the Advanced Trauma Life Support paradigm was available both as categorized variables and as free text. A non-exhaustive table of available pre-hospital variables was presented in the protocol of the GOAT study.[8]

In-Hospital Data

Evaluation of field triage accuracy requires information on resource use and injury severity of patients transported by EMSs to trauma-receiving emergency departments. Relevant patient outcomes collected by the LTR were, amongst others, Injury Severity Scores (ISS), early critical-resource use (e.g., damage control surgery), admission to the Intensive Care Unit, mortality, and hospital length of stay.[14]

The ISS is an anatomical scoring system used to classify injury severity. The American College of Surgeons Committee on Trauma suggested to use an ISS of 16 or greater as a reference standard for trauma system evaluation.[1] This implies that patients with an ISS of 16 or greater are considered to be severely injured and should be transported to higher-level trauma centers. The LTR calculates ISSs on the basis of the 2008 version of the Abbreviated Injury Scale.[15] Injury codes were calculated by trained trauma registrars after the final diagnosis was made.

Early critical-resource use is increasingly proposed as an alternative reference standard to assess the need for specialized trauma care.[16-18] Data on early critical-resource use were categorized into 11 categories, including damage control laparotomy/thoracotomy, craniotomy, and radiological interventions.

Health, Medical Consumption and Productivity Questionnaires

The TESLA trial introduced active follow-up at four timepoints. The EQ-5D-5L, the Short Form 36, and a subset of the Medical Consumption Questionnaire and the Productivity Cost Questionnaire were chosen to evaluate the impact of the new triage algorithms by means of a cost-utility analysis.[19-21] These questionnaires were chosen based on their expected validity in the current cohort which is very heterogeneous because of the inclusion of patients with all types of injury patterns and varying injury severity.

Record Linkage

A combined deterministic and probabilistic approach to patient linkage was applied to match EMS patient records to patient data from the LTR. Records were deterministically linked whenever a unique ambulance patient-record-identifier was available in both pre-hospital and in-hospital data.

A probabilistic software utility (LinkAssist) was developed to link anonymized patient records that lacked unique identifiers. Data from 28,634 deterministic matches were used to generate a training set (n=25,771) and a test set (n=2863) to develop and validate a robust prediction model specifically developed for accurate record linkage in TRACCC.

First, various predictors were selected that were available in both pre-hospital and in-hospital data (e.g., date of injury). Second, a training set was constructed by comparing each EMS record to a random sample of 30 entries from the LTR with a matching hospital and year of injury (including the actual matching record). Pre-hospital and hospital variables were compared using various similarity methods dependent on the type of the predictor. Third, a validation dataset was constructed from the hold-out deterministic matches following the same methodology, except for record subsampling in order to preserve the original matching versus non-matching ratio.

A gradient boosting machine was fit to the training data with hyperparameters that were optimized using 10-fold cross-validation. The predicted class was generated for each pair of EMS record and LTR record and evaluated against the true class (i.e., match or no match). The model proved to be accurate with an overall accuracy of 100% (95%-CI, 100 – 100), a sensitivity of 100% (95%-CI, 99 – 100), and specificity of 100% (95%-CI, 100 – 100). The c-statistic of the newly developed model was 1.0 (95%-CI, 1.0 – 1.0) in the test set. Binomial confidence intervals were calculated using the Agresti-Coull formula for each accuracy metric and the bootstrap (r=2000) was used to calculate confidence intervals for the c-statistic.

WHAT HAS IT FOUND? KEY FINDINGS AND PUBLICATIONS

TRACCC includes data from eight different EMSs in the Netherlands. Demographic characteristics of the included regions are displayed in Table 1. The participating EMSs are integrated into six inclusive trauma systems and their service area spans over 8000 km^2 with roughly 6.5 million inhabitants. Thirty-four trauma centers lie within these service areas, of which six level-I trauma centers and 28 level-II or III trauma centers.

Table 1 Demographic Characteristics of Regions Included in the Trauma Continuum of Care Cohort

	EMS region								
Statistic	**All regions**	**Rotterdam-Rijnmond**	**Amsterdam-Amstelland**	**Brabant Midden-West**	**Brabant Noord**	**Zaanstreek-Waterland**	**Gelderland Zuid**	**Utrecht**	**Zuid-Holland Zuid**
General									
Area, km^2	8096	866	282	2123	1333	348	1039	1385	720
Inhabitants, *n*	6,585,646	1,267,100	981,095	1,100,840	621,357	325,320	538,250	1,268,489	483,195
Hospitals*									
Total, *n*	34	7	6	5	3	2	3	5	3
Level-I, *n*	6	1	2	1	0	0	1	1	0
Level-II/III, *n*	28	6	4	4	3	2	2	4	3

Abbreviations: EMS, *Emergency Medical Service*.

Baseline characteristics of included patients are displayed in Table 2. The median age in the cohort was 57.1 years (IQR, 30.2 – 77.9). Nearly half of the included patients were women. The median response time (i.e., time from ambulance dispatch to arrival at the scene) was 8.8 minutes (IQR, 6.0 – 12.3). Approximately half of all urgently transported patients were transported with lights and sirens. The median systolic blood pressure was 139 mm Hg (IQR, 122 – 156) and 2644 patients had a systolic blood pressure <90 mm Hg. Respiratory rates were stable across all regions with a rate of 16 (IQR, 14 – 18) and 4624 (3%) had a respiratory rate of less than 10 or higher than 29. Nearly ten thousand patients (6%) had an impaired Glasgow Coma Scale.

One in three patients was hospitalized and median length of stay was 3.1 days (IQR, 1.0 – 7.2). The overall median ISS of hospitalized patients was 6, ranging from 5 to 9 between participating EMSs. A total of 3760 patients had an ISS of 16 or greater. Severe injuries were mostly present in the extremities (31%) followed by head and neck injuries (9%) and thoracic injuries (6%). A total of 3875 (8%) patients were in need of specialized trauma care according to a resource-based reference standard consisting of: discharge from the emergency department to the intensive care unit, out-of-hospital intubation, major interventions within 12 hours, and mortality <24 h.

The main result from the TRACCC cohort is the triage accuracy based on the initial transportation destination and three different reference standards: early critical-resource use, an anatomical reference standard (ISS ≥16), and ISS of 16 or greater on a subset of the cohort excluding patients with an ISS between 9 and 15 (i.e., patients that are considered not to be severely injured nor to be mildly injured). The overall undertriage rates in the current cohort were 24% (95%-CI, 22 – 25) for the anatomical reference standards (ISS ≥16) and 28% (95%-CI, 27 – 30) for the resource-based reference standard. Overtriage rates were 21% (95%-CI, 21 – 21) for the anatomical reference standard, 21% (95%-CI, 21 – 21) while excluding ISS 9–14, and 21% (95%-CI, 21 – 21). Sensitivity and specificity metrics are displayed in Table 3, alongside the values needed to reconstruct the contingency tables.

Research conducted on data from predecessors of TRACCC showed that triage accuracy based on the initial transportation destination did not comply with the guidelines of the ACSCOT in a single region.[1,9] A new tool to accurately identify trauma patients in the pre-hospital setting was created in response to this finding.[12] Other research investigated the diagnostic value of EMS professional judgment in the identification of head injuries and thoracic injuries, and the compliance to contemporary triage tools.[22-24]

Table 2 Baseline Characteristics of the Trauma Continuum of Care Cohort

Variable	EMS region								
	All patients (n=165,404)	**Rotterdam-Rijnmond (n=34,919)**	**Amsterdam-Amstelland (n=34,343)**	**Brabant Midden-West (n=30,194)**	**Brabant Noord (n=15,682)**	**Zaanstreek-Waterland (n=11,020)**	**Gelderland Zuid (n=11,949)**	**Utrecht (n=21,715)**	**Zuid-Holland Zuid (n=5582)**
General									
Age, y	57.1 (30.2 - 77.9)	56.6 (29.8 - 78.2)	52.3 (28.2 - 74.7)	58.4 (31.4 - 78.3)	61.0 (34.7 - 79.2)	59.5 (31.3 - 79.7)	59.5 (31.9 - 78.4)	57.7 (29.1 - 78.5)	61.1 (34.5 - 78.8)
Men	83,607 (50.5)	17,762 (50.9)	17,894 (52.1)	15,507 (51.4)	7893 (50.3)	5307 (48.2)	5985 (50.1)	10,514 (48.4)	2745 (49.2)
Women	81,797 (49.5)	17,157 (49.1)	16,449 (47.9)	14,687 (48.6)	7789 (49.7)	5713 (51.8)	5964 (49.9)	11,201 (51.6)	2837 (50.8)
Medical doctor on-scene	3032 (6.1)	759 (8.4)	681 (8.4)	524 (5.4)	288 (5.6)	246 (7.7)	251 (5.4)	169 (2.3)	114 (5.1)
Ambulance response time	8.8 (6.0 - 12.3)	8.0 (5.5 - 11.4)	7.7 (5.5 - 10.9)	9.4 (6.7 - 12.7)	9.7 (6.8 - 13.0)	8.0 (5.6 - 11.1)	9.8 (6.8 - 13.5)	10.0 (7.0 - 14.0)	8.8 (6.0 - 12.3)
Highest priority	25,253 (51.0)	4994 (55.1)	5755 (71.2)	4285 (44.0)	2254 (43.7)	2144 (67.1)	1756 (37.9)	2877 (38.9)	1188 (53.4)
Vital signs									
Systolic blood pressure	139 (122 - 156)	139 (122 - 157)	135 (120 - 152)	139 (123 - 156)	139 (124 - 156)	138 (121 - 157)	139 (122 - 158)	140 (124 - 158)	140 (125 - 160)
Systolic blood pressure <90	2644 (1.6)	543 (1.6)	619 (1.8)	425 (1.4)	242 (1.5)	196 (1.8)	193 (1.6)	344 (1.6)	83 (1.5)
Respiratory rate	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)
Respiratory rate <10 or >29	4624 (2.8)	1006 (2.9)	1083 (3.2)	669 (2.2)	408 (2.6)	362 (3.3)	360 (3.0)	606 (2.8)	130 (2.3)
Heart rate	83 (74 - 95)	84 (75 - 96)	84 (75 - 96)	82 (73 - 95)	81 (72 - 94)	84 (75 - 95)	80 (73 - 92)	80 (72 - 92)	83 (74 - 95)
Oxygen saturation	98 (96 - 99)	98 (96 - 99)	98 (96 - 99)	97 (96 - 99)	97 (96 - 98)	98 (96 - 98)	98 (96 - 99)	98 (96 - 99)	98 (96 - 99)
Glasgow Coma Scale <14	9651 (5.8)	1979 (5.7)	1915 (5.6)	1558 (5.2)	836 (5.3)	624 (5.7)	696 (5.8)	1594 (7.3)	449 (8.0)

Table 2 Continued

Variable	All patients (n=165,404)	EMS region							
		Rotterdam-Rijnmond (n=34,919)	Amsterdam-Amstelland (n=34,343)	Brabant Midden-West (n=30,194)	Brabant Noord (n=15,682)	Zaanstreek-Waterland (n=11,020)	Gelderland Zuid (n=11,949)	Utrecht (n=21,715)	Zuid-Holland Zuid (n=5582)
Outcomes									
In-hospital stay	49,472 (29.9)	9055 (25.9)	8078 (23.5)	9745 (32.3)	5157 (32.9)	3193 (29.0)	4627 (38.7)	7390 (34.0)	2227 (39.9)
Length of hospitalization*	3.1 (1.0 - 7.2)	3.1 (0.9 - 7.0)	3.0 (1.0 - 7.0)	3.2 (1.2 - 8.0)	3.2 (1.2 - 8.0)	3.0 (1.0 - 8.0)	3.0 (1.0 - 7.0)	4.0 (1.2 - 8.0)	2.7 (0.9 - 5.8)
ISS*	6 (4 - 9)	6 (4 - 9)	5 (4 - 9)	6 (3 - 9)	8 (4 - 9)	5 (2 - 9)	5 (4 - 9)	9 (4 - 9)	8 (4 - 9)
ISS ≥16*	3760 (7.6)	731 (8.1)	573 (7.1)	719 (7.4)	353 (6.8)	166 (5.2)	401 (8.7)	649 (8.8)	168 (7.5)
ISS score ≥3, per region*	22,550 (45.6)	4133 (45.6)	3441 (42.6)	4262 (43.7)	2422 (47.0)	1332 (41.7)	2026 (43.8)	3894 (52.7)	1040 (46.7)
Head and neck	4505 (9.1)	886 (9.8)	821 (10.2)	767 (7.9)	413 (8.0)	229 (7.2)	368 (8.0)	742 (10.0)	279 (12.5)
Face	81 (0.2)	23 (0.3)	4 (0.0)	17 (0.2)	8 (0.2)	3 (0.1)	7 (0.2)	16 (0.2)	3 (0.1)
Thorax	3066 (6.2)	588 (6.5)	447 (5.5)	520 (5.3)	324 (6.3)	169 (5.3)	298 (6.4)	581 (7.9)	139 (6.2)
Abdomen	674 (1.4)	153 (1.7)	94 (1.2)	114 (1.2)	63 (1.2)	25 (0.8)	79 (1.7)	114 (1.5)	32 (1.4)
Extremities	15,475 (31.3)	2727 (30.1)	2231 (27.6)	3076 (31.6)	1772 (34.4)	952 (29.8)	1403 (30.3)	2661 (36.0)	653 (29.3)
External	350 (0.7)	77 (0.9)	63 (0.8)	53 (0.5)	16 (0.3)	26 (0.8)	36 (0.8)	68 (0.9)	11 (0.5)
Early critical-resource use*	3875 (7.8)	812 (9.0)	550 (6.8)	714 (7.3)	336 (6.5)	180 (5.6)	408 (8.8)	727 (9.8)	148 (6.6)
Discharge from emergency department to ICU	2812 (5.7)	511 (5.6)	387 (4.8)	564 (5.8)	263 (5.1)	128 (4.0)	315 (6.8)	552 (7.5)	92 (4.1)
Out-of-hospital intubation	1141 (2.3)	236 (2.6)	185 (2.3)	149 (1.5)	113 (2.2)	82 (2.6)	118 (2.6)	193 (2.6)	65 (2.9)
Major intervention <12 h	906 (1.8)	299 (3.3)	146 (1.8)	135 (1.4)	55 (1.1)	33 (1.0)	74 (1.6)	121 (1.6)	43 (1.9)
Mortality <24 h	330 (0.7)	69 (0.8)	32 (0.4)	87 (0.9	24 (0.5)	14 (0.4)	25 (0.5)	63 (0.9)	16 (0.7)

Abbreviations: EMS, *Emergency Medical Service*; ICU, *Intensive Care* Unit; ISS, *Injury Severity Score*. Variables with missing values were multiply imputed. Data are represented as median, IQR or n (%) unless otherwise specified. *Hospitalized patients only.

Table 3 Overall Triage Accuracy in Participating EMS Regions

	Outcomes		
Variable	**ISS ≥16**	**ISS ≥16 (excl. ISS 9–14)**	**Early critical-resource use**
Transported to higher-level trauma center			
With outcome (TP)	2867	2867	3153
Without outcome (FP)	34,296	29,581	34,010
Transported to lower-level trauma center			
With outcome (FN)	893	893	1251
Without outcome (TN)	127,348	112,196	126,990
Undertriage	23.8 (22.4 - 25.1)	23.8 (22.4 - 25.1)	28.4 (27.1 - 29.8)
Overtriage	21.2 (21.0 - 21.4)	20.9 (20.7 - 21.1)	21.1 (20.9 - 21.3)
Sensitivity	76.3 (74.9 - 77.6)	76.3 (74.9 - 77.6)	71.6 (70.2 - 72.9)
Specificity	78.8 (78.6 - 79.0)	79.1 (78.9 - 79.3)	78.9 (78.7 - 79.1)

Abbreviations: EMS, *Emergency Medical Service*; FN, *false negative*; FP, *false positive*; ISS, *Injury Severity Score*; TN, *true negative*; TP, *true positive*. Values are presented as n or percentage (95%-CI).

WHAT ARE THE MAIN STRENGTHS AND WEAKNESSES?

Pre-hospital trauma care and triage accuracy in particular is highly dependent on the study population, varying injury patterns, and geographical differences (e.g., trauma center proximity) between trauma systems. The heterogeneous nature of trauma systems and EMSs limits the generalizability of diagnostic or prognostic studies conducted in a single region. In contrast, registry-based studies conducted in multiple trauma systems often suffer from a considerable amount of selection or information bias. TRACCC was designed to overcome these hindrances and enables researchers to prospectively study the effects of pre-hospital triage, treatment, and transportation on in-hospital and longitudinal patient outcomes on a large scale.

Another major strength of TRACCC is its computer-aided patient selection strategy that is employed to prevent selection bias. Selection bias is often present in triage research that involves patients that were selected based on the presumed chief complaint as recorded by EMS professionals. This variable was missing in 29% of all hand-labelled cases in TRACCC. In a subgroup of hand-labeled cases, the incidence of false-positively identified trauma patients was 5%, while 24% of the trauma patients had a label that

belonged to a different medical specialty such as general surgery, neurology, or pediatrics. Complete-case analyses or simple imputation strategies would not have been sufficient. In contrast, our computer-assisted patient selection proved to be highly accurate in an independent hold-out validation set.

The combined deterministic and probabilistic linkage methodology was exclusively developed to link EMS patient records with records from the LTR in TRACCC. The resulting record linkage strategy proved to be both highly accurate and computationally feasible. Automated record linkage allowed TRACCC to grow to a big volume and enabled new data to be processed instantaneously, while preserving data quality.

Finally, TRACCC allows researchers to study the accuracy of triage based on alternative reference standards such as a composite endpoint consisting of early critical-resource use, intensive care unit admission, pre-hospital intubation, and mortality. Further research on the data enclosed in TRACCC might allow researchers to establish a resource-based approach to trauma triage, which could potentially overcome limitations that are inherent to the use of surrogate markers of resource need, such as the ISS.

TRACCC also has certain limitations. Missing data is empirical in acute care settings in general and during field triage in particular. Vital signs were frequently not measured, such as systolic blood pressure (27%), respiratory rate (36%), heart rate (18%), oxygen saturation (29%) and Glasgow Coma Scale (14%). A multi-level multiple imputation strategy that accounted for clustering was employed to impute missing data in 48 datasets (12 per processor core). Although the proportion of missing data was substantial for certain variables, it does serve as an adequate representation of the pre-hospital setting. This setting is characterized by limited diagnostic modalities and a lack of complete information on which EMS professionals need to act.

CAN I GET HOLD OF THE DATA? WHERE CAN I FIND OUT MORE?

TRACCC was designed to facilitate researchers to study effects that transcend the pre-hospital setting. Collaboration in data analysis, publication, and new research proposals on this subject are welcomed. Researchers can submit research protocols to the scientific committee, which can be contacted through the web form available on https://www.traumatriageapp.com/en/. Details about the application process can be requested by contacting R.D. Lokerman [r.d.lokerman@umcutrecht.nl] or J.F. Waalwijk [job.waalwijk@nazl.nl].

Funding
This work was supported by the Netherlands Organization for Health Research and Development [grant number 80-84300-98-18555].

REFERENCES

1. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Aug 20, 2019).

2. Curtis K, Chong S, Mitchell R, Newcombe M, Black D, Langcake M. **Outcomes of severely injured adult trauma patients in an Australian health service: does trauma center level make a difference?** World J Surg 2011; 35(10): 2332-40.

3. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

4. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

5. van der Sluijs R, van Rein EAJ, Wijnand JGJ, Leenen LPH, van Heijl M. **Accuracy of pediatric trauma field triage: a systematic review.** JAMA Surg 2018; 153(7): 671-6.

6. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld RA, Leenen LPH, van Heijl M. **Accuracy of prehospital triage protocols in selecting severely injured patients: A systematic review.** J Trauma Acute Care Surg 2017; 83(2): 328-39.

7. van Rein EAJ, van der Sluijs R, Houwert RM, et al. **Effectiveness of prehospital trauma triage systems in selecting severely injured patients: Is comparative analysis possible?** Am J Emerg Med 2018; 36(6): 1060-9.

8. van der Sluijs R, Debray TPA, Poeze M, Leenen LPH, van Heijl M. **Development and validation of a novel prediction model to identify patients in need of specialized trauma care during field triage: design and rationale of the GOAT study.** Diagn Progn Res 2019; 3: 12.

9. Voskens FJ, van Rein EAJ, van der Sluijs R, et al. **Accuracy of prehospital triage in selecting severely injured trauma patients.** JAMA Surg 2018; 153(4): 322-7.

10. Kingma D, Ba J. **Adam: A Method for Stochastic Optimization.** Proceedings of the 3rd International Conference on Learning Representations 2014.

11. van der Sluijs R, Fiddelers AAA, Waalwijk JF, et al. **The impact of the Trauma Triage App on prehospital trauma triage: design of the stepped-wedge, cluster randomized TESLA trial [NTR7243]**. https://www.trialregister.nl/trial/7038 (accessed December 9 2019).

12. van Rein EAJ, van der Sluijs R, Voskens FJ, et al. **Development and validation of a prediction model for prehospital triage of trauma patients.** JAMA Surg 2019.

13. Ringdal KG, Coats TJ, Lefering R, et al. **The Utstein template for uniform reporting of data following major trauma: a joint revision by SCANTEM, TARN, DGU-TR and RITG.** Scand J Trauma Resusc Emerg Med 2008; 16: 7.

14. 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(3): 187-96.

15. Gennarelli T, Wodzin E. **The Abbreviated Injury Scale 2005: update 2008**. Association for the Advancement of Automotive Medicine; Des Plaines, IL: 2008.

16. Baxt WG, Jones G, Fortlage D. **The trauma triage rule: a new, resource-based approach to the prehospital identification of major trauma victims.** Ann Emerg Med 1990; 19(12): 1401-6.

17. 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.

18. Newgard CD, Hedges JR, Diggs B, Mullins RJ. **Establishing the need for trauma center care: anatomic injury or resource use?** Prehosp Emerg Care 2008; 12(4): 451-8.

19. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L. **The iMTA Productivity Cost Questionnaire: A standardized instrument for measuring and valuing health-related productivity losses.** Value Health 2015; 18(6): 753-8.

20. Ware JE, Jr., Sherbourne CD. **The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.** Med Care 1992; 30(6): 473-83.

21. Janssen MF, Pickard AS, Golicki D, et al. **Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study.** Qual Life Res 2013; 22(7): 1717-27.

22. van Rein EAJ, Jochems D, Lokerman RD, et al. **Diagnostic value of emergency medical services provider judgement in the identification of head injuries among trauma patients.** Eur J Neurol 2019; 26(2): 274-80.

23. van Rein EAJ, Lokerman RD, van der Sluijs R, et al. **Identification of thoracic injuries by emergency medical services providers among trauma patients.** Injury 2019; 50(5): 1036-41.

24. van Rein EAJ, Sadiqi S, Lansink KWW, et al. **The role of emergency medical service providers in the decision-making process of prehospital trauma triage.** Eur J Trauma Emerg Surg 2018.

PART III

Personalized field triage

CHAPTER VI

Development and validation of a prediction model for pre-hospital triage of trauma patients

JAMA Surgery

Eveline A.J. van Rein; Rogier van der Sluijs; Frank J. Voskens; Koen W.W. Lansink; R. Marijn Houwert; Rob A. Lichtveld; Mariska A. de Jongh; Marcel G.W. Dijkgraaf; Howard R. Champion; Frank J.P. Beeres; Luke P.H. Leenen; Mark van Heijl

ABSTRACT

Importance

Pre-hospital trauma triage protocols are used worldwide to get the right patient to the right hospital and thereby improve the chance of survival and avert lifelong disabilities. The American College of Surgeons Committee on Trauma set target levels for undertriage rates of less than 5%. None of the existing triage protocols has been able to achieve this target in isolation.

Objective

To develop and validate a new pre-hospital trauma triage protocol to improve current triage rates.

Design, Setting, and Participants

In this multicenter cohort study, all patients with trauma who were 16 years and older and transported to a trauma center in two different regions of the Netherlands were included in the analysis. Data were collected from Jan 1, 2012, through Jun 30, 2014, in the Central-Netherlands region for the design data cohort and from Jan 1 through Dec 31, 2015, in the Brabant region for the validation cohort. Data were analyzed from May 3, 2017, through Jul 19, 2018.

Main Outcomes and Measures

A new prediction model was developed in the Central-Netherlands region based on pre-hospital predictors associated with severe injury. Severe injury was defined as an Injury Severity Score greater than 15. A full-model strategy with penalized maximum likelihood estimation was used to construct a model with eight predictors that were chosen based on clinical reasoning. Accuracy of the developed prediction model was assessed in terms of discrimination and calibration. The model was externally validated in the Brabant region.

Results

Using data from 4950 patients with trauma from the Central-Netherlands region for the design data set (58.3% male; mean [SD] age, 47 [21] years) and 6859 patients for the validation Brabant region (52.2% male; mean [SD] age, 51 [22] years), the following eight significant predictors were selected for the prediction model: age; systolic blood pressure; Glasgow Coma Scale score; mechanism criteria; penetrating injury to the head, thorax, or abdomen; signs and/or symptoms of head or neck injury; expected injury in the Abbreviated Injury Scale thorax region; and expected injury in two or more Abbreviated Injury Scale regions. The prediction model showed a c-statistic of 0.823 (95% CI, 0.813-

0.832) and good calibration. The cut-off point with a minimum specificity of 50.0% (95% CI, 49.3%-50.7%) led to a sensitivity of 88.8% (95% CI, 87.5%-90.0%). External validation showed a c-statistic of 0.831 (95% CI, 0.814-0.848) and adequate calibration.

Conclusions and Relevance
The new pre-hospital trauma triage prediction model may lower undertriage rates to approximately 10% with an overtriage rate of 50%. The next step should be to implement this prediction model with the use of a mobile app for Emergency Medical Services professionals.

INTRODUCTION

In first world countries, systems of trauma care substantially reduce mortality associated with injury.[1-3] Multiple studies focused on optimizing such trauma systems by balancing timely access to expert care, the ability of practitioners and teams to attain and sustain the necessary expertise, and the cost-effectiveness of the overall trauma system.[2,3] Fundamental to the trauma system is pre-hospital trauma triage, the goal of which is to identify at-risk patients and provide early and resuscitative care while transporting the patient to the highest appropriate level of care.[4] Identification of severely injured patients is challenging. Only 0.5% of those injured are severely injured.[4] Other challenges include assessment of the incident scene and the patient's physiologic state and risk of deterioration, identification of obvious injuries, and consideration of adjuvant factors, such as age.

Emergency Medical Services (EMS) professionals must differentiate between patients at the scene, often in adverse situations, without use of advanced diagnostic tools. Therefore, the EMS professionals are often forced to perform the pre-hospital trauma triage based on incomplete data. The importance of pre-hospital trauma triage cannot be understated; a structured and reliable process is crucial.

Worldwide, protocols are used to help identify severely injured patients. However, none of the existing protocols can achieve the recommended triage rates.[5-7] All are simplistic and static tools, whereas patients are dynamic, and more advanced methods are available to use at the scene, such as a prediction model in a mobile app.[8,9]

Undertriage occurs when severely injured patients are not transported to a higher-level trauma center and results in delayed care and increased mortality and morbidity.[1,2] Overtriage occurs when patients without severe injuries are taken to a higher-level trauma

center, often incurring preventable cost and resource consumption.[3,10] In the Netherlands, level-I trauma centers are considered higher-level trauma centers and level-II and III are considered lower-level trauma centers. The American College of Surgeons Committee on Trauma set target levels of undertriage at less than 5% and an overtriage as great as 25% to 35%.[11] The National Health Institute of the Netherlands recommends an undertriage rate of less than 10%. A target for the overtriage rate was not set; the overtriage rate, however, should depend on the regional circumstances and could be as great as 50%.

In the Netherlands, a protocol based on the Field Triage Decision Scheme, is used nationwide (see Appendix).[12] The mean undertriage rate was 33% in 2015.[13] A recent study showed an undertriage rate of 22% and an overtriage rate of 31% in a single inclusive trauma region.[14] At present, existing protocols achieve undertriage rates ranging from 22% to 44%, with overtriage rates ranging from 11% to 22% in a general trauma population.[6,15,16] The aim of the present study was therefore to develop and validate a new pre-hospital trauma triage prediction model that attempts to lower the undertriage rate to approximately 10%, with a maximum overtriage rate of 50%.

METHODS

Study Design and Setting

In the Netherlands, each ambulance service serves a region. In this prospective, multi-center cohort study, all adult patients with trauma and all trauma centers in two different regions were included. Data from the Central-Netherlands region were used to develop a diagnostic prediction model that was externally validated using the data of the Brabant region.
Central-Netherlands has one level-I trauma center, the University Medical Center Utrecht, and seven level-II or III trauma centers. The region covers 535 square miles and has 1.2 million residents. Brabant has one level-I trauma center, Elisabeth-TweeSteden Hospital Tilburg, and 10 level-II or III trauma centers. The region covers 1343 square miles and has 1.7 million residents. In both regions, the Dutch National Protocol of Ambulance Services is used (see Appendix).[12] Patients transported across the borders of these regions were excluded because data were unavailable.

Patients

All patients with trauma who were 16 years or older, determined to be highest priority (with flashing lights and sirens) by the dispatch center, and transported to a trauma center in one of the two regions underwent evaluation. In Central-Netherlands, patients were included between Jan 2012 and Jun 2014.[14] In Brabant, patients were included between Jan 2015 and Dec 2015.

Outcomes and Definitions

Independent predictors associated with severe injury were identified to create a prediction model consisting of a limited number of variables. A severely injured patient was defined as a patient with an Injury Severity Score (ISS) greater than 15 (range, 0–75).
Undertriage was defined as the proportion of severely injured patients transported to a level-II or III trauma center. Overtriage was defined as the proportion of patients without severe injuries transported to a level-I trauma center. The protocol allows for these patients to be transported to a level-I trauma center if it happens to be the nearest hospital.

Data Sources

Pre-hospital reports from EMS professionals were prospectively collected and included patient demographics, vital signs, description of the trauma mechanism, on-scene physical examination data (including suspected injured body region), and the receiving hospital. The body region suspected of injury by EMS professionals was divided into the head and neck, face, thorax, abdomen, extremities, or skin and others. These regions were chosen based on the categorization of the Abbreviated Injury Scale (AIS) regions that make up the ISS.

The Dutch National Trauma Database registered injuries for all patients admitted to a hospital. The Central-Netherlands patient data were also extracted from the electronic patient documentation for patients discharged directly from the emergency department. The injuries were recorded after discharge or 30 days after admission and coded by trained data managers in both regions (including E.A.J.v.R. and F.J.V.). Before 2015, all injuries were coded using the AIS 1990, Update 1998 (AIS98); from 2015 and on, the AIS 2005, Update 2008 (AIS08) was used. Therefore, the injuries were coded according to the AIS98 for the Central-Netherlands database and according to the AIS08 for the Brabant database. The ISS was calculated based on AIS scores to determine injury severity.

Missing Data

Missing data were analyzed and considered to be missing at random. Multiple imputation by chained equations was used for both regions separately to account for missing pre-hospital variables rather than deleting patients who had the most data available.[17] Missing values were imputed based on all other predictors as well as the ISS. For both regions, respiratory rate, systolic blood pressure, oxygen saturation level, and Glasgow Coma Scale score were imputed. Pulse was imputed for the Central-Netherlands region only because this variable was missing in the Brabant data. For the Brabant region, ISS was only available for admitted patients. An ISS less than 15 was assumed for patients discharged from the emergency department because a previous study demonstrated that all severely injured patients (ISS >15) were admitted or died in the emergency department.[14]

Statistical Analysis

Data were analyzed from May 3, 2017, through Jul 19, 2018. Frequencies with percentages were used to describe nominal and ordinal variables, whereas means and standard deviations were used to describe continuous variables. Bivariable binary logistic regression was used to explore potential predictors associated with severe injury (ISS >15). Analyses were performed on five imputed datasets independently and pooled using Rubin's rules, if applicable.[18]

To ensure practical applicability, the maximum number of predictors was limited to eight. A full-model strategy with clinically relevant variables was used to develop the prediction model. To improve the accuracy for future patients and other regions, penalized maximum likelihood estimation was used.[19] Penalized maximum likelihood estimation is a rigorous estimation method that potentially results in better generalizability, model reduction, and differential shrinkage of coefficients.[20] The functional forms of all continuous predictors were defined before modeling. Restricted cubic splines were used to model non-linear predictors.

The performance of the final prediction model was expressed in terms of discrimination and calibration. Discriminative value was quantified by the c-statistic. The receiver operating characteristic curve was plotted and a predefined value for specificity – an overtriage rate of 50% – was used to determine a cut-off point. Calibration was graphically assessed using a calibration plot.

The final model was externally validated with the Brabant data. Owing to heterogeneity between trauma regions (e.g., prevalence of severe injury), the model needed to be recalibrated by updating the intercept for the Brabant region.[21,22] Calibration and discrimination of the prediction model were assessed in the validation process. Bootstrapped 95% confidence intervals based on percentiles and 1000 resamples were calculated for c-statistics and accuracy metrics. Two-sided *P* values were calculated using univariable logistic regression, and $P < .05$ indicated significance. All statistical analyses were performed using R software (version 3.2.4, R Foundation for Statistical Computing).[23]

RESULTS

Design Data Set

A total of 4950 adult patients with trauma were included for the Central-Netherlands region constituting the design data set. To account for missing data, multiple imputation was used for respiratory rate in 324 (6.5%), systolic blood pressure in 345 (7.0%), oxygen saturation level in 662 (13.4%), and Glasgow Coma Scale score in 230 (4.6%). Mean patient age was 47 years (21) years; 2887 (58.3%) were male and 2063 (41.7%) were female; and 435 (8.8%) had an ISS greater than 15 (Table 1). In this cohort, the undertriage rate was 21.6%; the overtriage rate, 30.6%.

Table 1 Baseline Characteristics of the Central-Netherlands Region and the Brabant Region

Characteristic	Central-Netherlands region (n=4950)	Brabant region (n=6859)
Demographics		
Age, y	47 (21.3)	51 (22.1)
Men	2887 (58.3)	3583 (52.2)
Pregnancy	32 (0.6)	25 (0.4)
Use of oral anticoagulants	131 (2.6)	234 (3.4)
Alcohol use	531 (10.7)	746 (10.9)
Other drug use	43 (0.9)	39 (0.6)
Physiologic characteristics		
Systolic blood pressure, mm Hg	139 (23.6)	140 (24.3)
Respiratory rate, breaths/min	16 (4.0)	16 (5.1)
Oxygen saturation level	96 (4.3)	97 (3.1)
Glasgow Coma Scale score	14 (1.9)	15 (1.8)
Revised Trauma Score	12 (0.8)	NA
ABC unstable[a]	117 (2.7)	129 (1.9)
Mechanism of injury		
Mechanism criteria[b]	819 (16.5)	475 (6.9)
Fall 2-5 m Fall >5 m or >3 × body length	314 (6.3) 77 (1.6)	197 (2.9) 24 (0.3)
Fall from stairs, 1-10 steps Fall from stairs, >10 steps	388 (7.8) 86 (1.7)	288 (4.2) 87 (1.3)
Vehicle rollover	96 (1.9)	129 (1.9)

Characteristic	Central-Netherlands region (n=4950)	Brabant region (n=6859)
Injury characteristics		
Penetrating injury to head, thorax, or abdomen	90 (1.8)	30 (0.4)
Expected (unstable) pelvic fracture	26 (0.5)	11 (0.2)
Neurologic deficit (≥1 extremity)	75 (1.5)	60 (0.9)
Symptoms of cerebral contusion or concussion	348 (7.0)	516 (7.5)
Agitation	172 (3.5)	66 (1.0)
Expected injury in AIS region		
Head/neck or trauma to head	2635 (53.2)	2393 (34.9)
Face	955 (19.3)	977 (14.2)
Thorax	719 (14.5)	329 (4.8)
Abdomen	332 (6.7)	74 (1.1)
Extremities	2013 (40.7)	1501 (21.9)
Skin and others	85 (1.7)	85 (1.2)
Expected injury in ≥2 AIS regions	1230 (24.8)	309 (4.5)
Burning wound with or without inhalation trauma	77 (1.7)	80 (1.2)
Inhalation trauma	29 (0.6)	38 (0.6)
Clinical characteristics		
ISS	5 (7.1)	NA
ISS >15	435 (8.8)	165 (2.4)
Destination		
Level-I trauma center	1724 (34.8)	1882 (27.4)
Level-II trauma center	1326 (26.8)	4208 (61.4)
Level-III trauma center	1900 (38.4)	769 (11.2)
Admission to hospital	2047 (41.4)	1842 (26.9)
In-hospital death	61 (1.2)	57 (0.8)

Abbreviations: AIS, *Abbreviated Injury Scale*; ISS, *Injury Severity Score;* NA, *not applicable*. Data are represented as mean (SD) or n (%).

[a] Defined as systolic blood pressure less than 90 mm Hg and/or respiratory frequency greater than 29 breaths/min.

[b] Mechanism criteria include fall of greater than 2 m, motor vehicle crash at greater than 32 km/h, or any type of entrapment.

Validation Data Set

In the Brabant region, a total of 6859 adult trauma patients were included in the validation set. To account for missing data, multiple imputation was used for respiratory rate in 1973 patients (28.8%), systolic blood pressure in 1145 (16.7%), oxygen saturation level in 1431 (20.9%), and Glasgow Coma Scale score in 288 (4.2%). Mean patient age was 51 (22) years; 3583 (52.2%) were male and 3276 (47.8%) were female; and 165 (2.4%) had an ISS greater than 15 (Table 1). The ISS was only available for the admitted patients. In this cohort, the undertriage rate was 27.3%; the overtriage rate, 26.3%.

Model Development and Specification

To develop the prediction model, 43 potential pre-hospital predictors from the Central-Netherlands database were explored using bivariable analysis (Table 2). The following eight predictors were chosen for the final model, based on clinical reasoning: age; systolic blood pressure; Glasgow Coma Scale score; mechanism criteria; penetrating injury to the head, thorax, or abdomen; signs and/or symptoms of head or neck injury; expected injury in the AIS thorax region; and expected injury in two or more AIS regions. The optimal cut-off point with a minimum specificity of 50.0% (95%-CI, 49.3 – 50.7), led to a sensitivity of 88.8% (95%-CI, 87.5 – 90.0).

Model Performance

This prediction model resulted in an undertriage rate of 11.2% (Δ difference, 10.4%) and an overtriage rate of 50.0% (Δ difference, 19.4%) for the Central-Netherlands region. Robust estimation using penalized maximum likelihood showed that all variables in the model were significant independent predictors (Table 3). The model had a good discrimination with a c-statistic of 0.823 (95%-CI, 0.813 – 0.832). The recalibration method led to an intercept of 0.894 for the Brabant region. External validation using the Brabant region database showed that the model with the new intercept was well calibrated (Figure 1) and had a good discrimination with a c-statistic of 0.831 (95%-CI, 0.814 – 0.848).

Table 2 Bivariable Logistic Regression Analysis on the Central-Netherlands Region

Variables	β coefficient (SD)	*P* value	OR
Patient characteristics			
Age, y	0.010 (0.002)	<.001	1.010
Female	-0.447 (0.107)	<.001	0.640
Alcohol use	0.205 (0.152)	.18	1.228
Use of oral anticoagulants	0.047 (0.307)	.88	1.048
Physiologic characteristics			
Systolic blood pressure	0.005 (0.002)	.02	1.005
Systolic blood pressure <90 mmHg	0.818 (0.320)	.01	2.265
Pulse	0.009 (0.003)	.001	1.009
Respiratory rate	0.042 (0.011)	.001	1.043
Respiratory rate <10 or >29 /min	1.477 (0.238)	<.001	4.381
Oxygen saturation	-0.096 (0.009)	<.001	0.908
Glasgow Coma Scale score	-0.357 (0.019)	<.001	0.700
Revised Trauma Score	-0.846 (0.059)	<.001	0.429
ABC unstable[a]	2.209 (0.300)	<.001	9.110
Mechanism of injury			
Mechanism criteria[b]	1.272 (0.108)	<.001	3.566
Fall 2-5 m	0.628 (0.168)	.002	1.874
Fall >5 m or >3 times body length	1.777 (0.244)	<.001	5.910
Motor vehicle crash >65 km/h	-0.243 (0.200)	.22	0.784
Motorcycle crash >32 km/h	1.011 (0.150)	<.001	2.749
Vehicle deformity >50 cm	0.622 (0.488)	.20	1.863
Vehicle intrusion passenger compartment >30 cm	1.997 (0.495)	<.001	7.368
Vehicle rollover	0.073 (0.354)	.84	1.075
Motor vehicle vs pedestrian impact >10 km/h	0.599 (0.294	.04	1.820
Motor vehicle vs bicycle impact >10 km/h	0.382 (0.185)	.04	1.465
No helmet on motorcycle or horse	1.340 (0.243)	<.001	3.819
No seatbelt in vehicle in high-energy trauma	-0.247 (0.521	.64	0.781

Variables	β coefficient (SD)	*P* value	OR
Deployed airbag in motor vehicle crash	-0.559 (0.314)	.08	0.572
Entrapment in vehicle	1.328 (0.272)	<.001	3.773
Entrapment elsewhere	1.292 (0.370)	<.001	3.640
Trauma to the head	1.206 (0.109)	<.001	3.340
Suicide attempt	0.890 (0.324)	.006	2.435
Injury characteristics			
Penetrating injury to head, thorax, or abdomen	1.248 (0.251)	<.001	3.484
Expected (unstable) pelvic fracture	2.683 (0.400)	<.001	14.623
Neurologic deficit (≥1 extremity)	0.590 (0.330)	.07	1.804
Pupil difference	2.569 (0.300)	<.001	13.049
Symptoms of cerebral contusion or concussion	0.897 (0.150)	<.001	2.451
Agitation	1.782 (0.170)	<.001	5.939
Vomiting	0.920 (0.274)	.001	2.510
Signs and/or symptoms of head or neck injury	1.157 (0.117)	<.001	3.182
Expected injury in AIS region			
Face	0.412 (0.116)	<.001	1.510
Thorax	0.445 (0.127)	<.001	1.561
Abdomen	0.252 (0.184)	.17	1.286
Extremities	-0.190 (0.104)	.07	0.827
Expected injury in ≥2 AIS regions	1.110 (0.102)	<.001	3.035
Expected injury to spine	-0.344 (0.131)	.009	0.709

Abbreviations: AIS, *Abbreviated Injury Scale*; OR, *odds ratio*. Multiple imputation was used to account for the missing pre-hospital variables.

[a] Defined as systolic blood pressure less than 90 mm Hg and/or respiratory frequency greater than 29 breaths/min.

[b] Include fall of greater than 2 m, motor vehicle crash at a rate of greater than 32 km/h, or any type of entrapment.

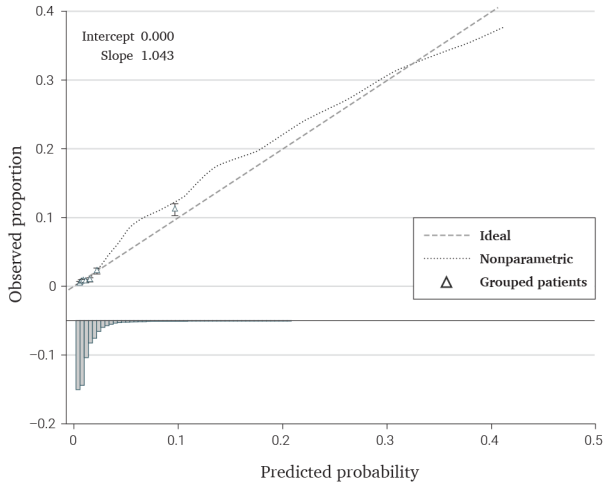


Figure 1 Calibration Plot of External Validation (Brabant Region)

DISCUSSION

In this prospective, multicenter cohort study, we present a ready-to-use pre-hospital trauma triage prediction model for the presence of severe injury at the scene in patients with trauma. The model performed well in the derivation set and in external validation. To our knowledge, this is the first externally validated protocol showing acceptable triage rates, with a potential undertriage rate of 11.2% and overtriage rate of 50.0%, depending on the chosen threshold.

Worldwide, triage protocols are based on a simple flowchart including vital signs, injury type, and mechanism of injury criteria.[6,24-27] These triage protocols are simplistic and static: transport to a higher-level trauma center should be considered if just one criterion is present. In reality, some factors have a greater association with injury severity than others, and the combination of factors indicates the need for higher-level trauma care. In addition, current protocols often use cut-off points for continuous variables, whereas the prediction model uses coefficients for each predictor to represent each variable's distinct association with injury severity to increase predictive ability.

The prediction model was based on three key elements: (1) inclusion of all adult trauma patients transported by an ambulance to (2) all trauma centers of an entire geographic region, with (3) pre-hospital parameters measured at the scene by EMS professionals. Previous studies have attempted to develop protocols but have not included these three key elements. For example, Dihn and colleagues[26] developed a triage protocol based on

Table 3. Multivariate Analysis of the Predictors for ISS >15 in the Central-Netherlands Region

Variables	β coefficient	*P* value	OR
Patient characteristics			
Age, y			
Spline basis function 1	0.011 (0.004)	.001	1.011
Spline basis function 2	0.001 (0.005)	.86	1.001
Physiologic characteristics			
Systolic blood pressure			
Spline basis function 1	-0.011 (0.002)	<.001	0.989
Spline basis function 2	0.020 (0.003)	<.001	1.020
Glasgow Coma Scale score	-0.337 (0.001)	<.001	0.714
Mechanism of injury			
Mechanism criteria[a]	1.314 (0.056)	<.001	3.721
Injury characteristics			
Penetrating injury to head, thorax, or abdomen	1.196 (0.131)	<.001	3.307
Signs and/or symptoms of head or neck injury	0.571 (0.056)	<.001	1.770
Expected injury in AIS region of thorax	0.405 (0.071)	<.001	1.499
Expected injury in ≥2 AIS regions	0.713 (0.129)	<.001	2.040
Intercept[b]	2.069 (0.315)	<.001	7.917

Abbreviations: AIS, *Abbreviated Injury Scale*; ISS, *Injury Severity Score*; OR, *odds ratio*. Multiple imputation was used to account for the missing pre-hospital variables.
[a] Include fall of greater than 2 m, motorcycle crash of greater than 32 km/h, or entrapment of a person or body part.
[b] Intercept is 0.894 for the Brabant region.

patients taken to a higher-level trauma center only, thereby excluding the undertriaged patients. Others included admitted patients only[28,29], thus excluding patients discharged from the emergency department or the potentially overtriaged patients. These models would not be reliable in a general trauma population because they fail to include the patient populations in which improvement is of utmost importance: the undertriaged and overtriaged patients.

Eight predictors were included in the prediction model based on clinical reasoning to achieve the best accuracy, while keeping it user friendly without too many factors. Age was included because previous studies showed a higher undertriage rate in elderly

patients.[6,30-32] Two continuous predictors of the condition of a patient are systolic blood pressure and Glasgow Coma Scale score.[33-35] Penetrating injury is an obvious predictor associated with potential severe trauma. The brain and thorax are two of the most commonly injured body regions, with both associated with a high prevalence of severe injury.[36-38] Also, multiregional injury was previously found as a strong predictor associated with severe injury.[26] Therefore, these eight predictors were included in the current prediction model. The prediction model resulted in an undertriage rate of 11.2% and an overtriage rate of 50.0% in Central-Netherlands.

After penalized estimation, the updated diagnostic prediction model was externally validated with data from the Brabant region. The prediction model requires an update primarily owing to the difference in prevalence of severe injury, resulting in a difference in baseline risk.[39] To account for this difference, the constant value (intercept) in the equation was altered. The constant value can be altered for other regions before applying the prediction model based on pre-hospital and hospital data of the specific region. External validation – with the altered intercept – showed good discrimination and calibration, indicating that the prediction model would likely be accurate in a region that is heterogeneous with respect to population, prevalence of severe injury, and mechanism of injury. Additionally, the injuries were coded differently in both regions: AIS98 was used in the Central-Netherlands region and AIS08 in the Brabant region. When using the AIS08, the overall ISS is lower compared with injuries scored with AIS98.[40] External validation using the AIS08 showed that the prediction model functions well using the most recent AIS.

The prediction model did not achieve the goal of an undertriage rate of less than 5% as targeted by the American College of Surgeons Committee on Trauma.[41] However, the model is a significant improvement compared with the existing protocols.[5,6,15,42] Whether further improvement of undertriage is achievable by solely improving the protocol is unclear. Previous studies have shown that addition of the judgment of EMS professionals can be useful in the identification of severely injured patients.[43-46] In this study, EMS professional judgment could not be quantified because it was not recorded. Including EMS professional judgment may improve the undertriage rate even more as well as increase adherence to the protocol. Other factors could have influenced transport decisions, such as geographical distance to a higher-level trauma center. Although distances are relatively small in the Netherlands compared with other countries, distance could have influenced the destination decision, especially in Brabant, because distances are larger in this region. Unfortunately, the effect of distance on the triage quality could not be evaluated in this study, and the distance at which deviation to the nearest hospital is best remains unclear.[47]

In practice, it is not feasible to calculate the risk of severe injury based on an equation that must be memorized and applied at the scene. This problem could be solved by implementing the prediction model in a mobile app; such triage tools are increasingly being developed and used in the pre-hospital process.[8,9] This mobile app includes every variable, calculates the chance of severe injury, and gives advice on where to transport the patient, which is much more practical for EMS professionals compared with an equation. With a mobile device available on every ambulance, the EMS professionals can calculate the risk of severe injury using the prediction model in the app to decide quickly and more accurately where to transport the patient. The EMS professional judgment could be included in the app. A mobile app with the described prediction model is currently being implemented in different regions in the Netherlands. The implementation aims at reducing undertriage rates specifically.

Limitations

This study has several limitations. In the final prediction model, missing data were present for systolic blood pressure and Glasgow Coma Scale variables. The data were considered to be missing at random, and multiple imputation was used to minimize selection bias. Second, for Brabant, the ISS was only available for the patients who were admitted or who died in the emergency department. In Central-Netherlands – where the ISS was available for all patients – all severely injured patients (ISS >15) were admitted or died in the emergency department. Accordingly, an ISS less than 16 was assumed for patients discharged from the emergency department. Last, debate remains on most accurate definition of a severely injured patient. Legitimate classification is difficult and dependent on multiple factors, such as regional circumstances and trauma center level. An ISS greater than 15 might not represent all patients in need of higher-level trauma center resources. However, an ISS greater than 15 is the most used surrogate marker for a severely injured patient when evaluating pre-hospital trauma triage; therefore, we chose this variable to define a severely injured patient.[5]

Future research should focus on the validation of the prediction model in other regions. Differences in prevalence of severely injured patients, and consequently baseline risk, can be large, and therefore different baseline risks should be determined in other populations. A possible solution is to validate the prediction model in other regions using our methodology.

The mobile app has been developed and is currently being implemented in the Netherlands (Figure 2). In this mobile app, the equation is integrated in addition to EMS professional judgment. This combination could be optimal to improve triage rates. Additionally, it could give insight in the value of EMS professional judgment.

For a penetrating injury to the head, thorax, or abdomen, the EMS professional can select yes or no.

The body regions where injury is expected can be selected. In this situation, the body region "thorax" is selected.

Baseline on EMS professional judgment, does the patient need to be transported to a level-I trauma center?

The advice of the TTApp is to transport the patient to a level-I trauma center.

Figure 2 Screenshots of the Mobile App. Abbreviations: EMS, *Emergency Medical Service*.

CONCLUSIONS

To our knowledge, this is the first study to develop and validate a pre-hospital trauma triage protocol based on all adult patients with trauma to be transported to a trauma center within a region and that can lower the undertriage rate to 11.2%, with an overtriage rate of 50.0%, from 21.6% and 30.6%, respectively. This protocol, based on an equation in which each predictor has its own coefficient, can be implemented with a mobile app for EMS professionals and could be of great help to lower undertriage rates.

REFERENCES

1. MacKenzie EJ, Rivara FP, Jurkovich GJ et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354:366-378.

2. Staudenmayer K, Weiser TG, Maggio PM, Spain DA, Hsia RY. **Trauma center care is associated with reduced readmissions after injury.** J Trauma Acute Care Surg 2016; 80:412-416.

3. Newgard CD, Staudenmayer K, Hsia RY et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32:1591-1599.

4. Champion HR, Lombardo LV, Wade CE, Kalin EJ, Lawnick MM, Holcomb JB. **Time and place of death from automobile crashes: Research endpoint implications.** J Trauma Acute Care Surg 2016; 81:420-426.

5. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld R, Leenen LPH, van Heijl M. **Accuracy of pre-hospital triage protocols in selecting major trauma patients: a systematic review.** J Trauma Acute Care Surg 2017; 83:328-339.

6. Newgard CD, Fu R, Zive D et al. **Prospective validation of the national field triage guidelines for identifying seriously injured persons.** J Am Coll Surg 2016; 222:146-158.

7. Newgard CD, Zive D, Holmes JF et al. **A multisite assessment of the American College of Surgeons Committee on Trauma field triage decision scheme for identifying seriously injured children and adults.** J Am Coll Surg 2011; 213:709-721.

8. Lewis TL, Fothergill RT, Aneurysm-FILTR Study Group, Karthikesalingam A. **Ambulance smartphone tool for field triage of ruptured aortic aneurysms (FILTR): study protocol for a prospective observational validation of diagnostic accuracy.** BMJ Open 2016;6:e011308.

9. Nogueira RG, Silva GS, Lima FO et al. **The FAST-ED App: A smartphone platform for the field triage of patients with stroke.** Stroke 2017;48.

10. Faul M, Xu L, Sasser SM. **Hospitalized traumatic brain injury: low trauma center utilization and high interfacility transfers among older adults.** Prehosp Emerg Care 2016; 20:594-600.

11. Sasser SM, Hunt RC, Faul M et al. **Guidelines for Field Triage of Injured Patients: Recommendations of the National Expert Panel on Field Triage, 2011.** MMWR Recomm Rep 2012; 61:1-20.

12. Ambulancezorg Nederland. **Landelijk Protocol Ambulancezorg (versie 7.1).** Zwolle, 2007. Stichting Landelijke Ambulance en Meldkamer Protocollen (LAMP).

13. Landelijk Netwerk Acute Zorg. **Landelijke Traumaregistratie 2011-2015: Rapportage Nederland.** Utrecht, 2016.

14. Voskens FJ, van Rein EAJ, van der Sluijs R, et al. **Accuracy of prehospital triage in selecting severely injured trauma patients.** JAMA Surg 2018; 153(4): 322-7.

15. Bond RJ, Kortbeek JB, Preshaw RM. **Field trauma triage: combining mechanism of injury with the Pre-hospital Index for an improved trauma triage tool.** J Trauma 1997; 43:283-287.

16. Meisler R, Thomsen AB, Abildstrøm H et al. **Triage and mortality in 2875 consecutive trauma patients.** Acta Anaesthesiol Scand 2010; 54:218-223.

17. van Buuren S, Groothuis-Oudshoorn K. **mice: Multivariate Imputation by Chained Equations in R.** J Stat Softw 2011;45.

18. Marshall A, Altman DG, Holder RL, Royston R. **Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines.** BMC Med Res Methodol 2009; 9:57-64.

19. Harrell FE. **Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis.** New York: Springer, 2001.

20. Moons KGM, Donders ART, Steyerberg EW, Harrell FE. **Penalized maximum likelihood estimation to directly adjust diagnostic and prognostic prediction models for overoptimism: a clinical example.** J Clin Epidemiol 2004; 57:1262- 1270.

21. Vergouwe Y, Nieboer D, Oostenbrink R et al. **A closed testing procedure to select an appropriate method for updating prediction models.** Stat Med 2016.

22. Janssen KJM, Moons KGM, Kalkman CJ, Grobbee DE, Vergouwe Y. **Updating methods improved the performance of a clinical prediction model in new patients.** J Clin Epidemiol 2008; 61:76-86.

23. R Development Core Team. **R: A language and environment for statistical computing.** Vienna, Austria, 2016. R Foundation for Statistical Computing.

24. van Laarhoven JJ, Lansink KW, van Heijl M, Lichtveld RA, Leenen LP. **Accuracy of the field triage protocol in selecting severely injured patients after high energy trauma.** Injury 2014; 45:869-873.

25. Hamada SR, Gauss T, Duchateau FX et al. **Evaluation of the performance of French physician-staffed emergency medical service in the triage of major trauma patients.** J Trauma Acute Care Surg 2014; 76:1476-1483.

26. Dihn MM, Bein KJ, Oliver M, Veillard AS, Ivers R. **Refining the trauma triage algorithm at an Australian major trauma centre: derivation and internal validation of a triage risk score.** Eur J Trauma Emerg Surg 2014;40:67-74.

27. Rubenson Wahlin R, Ponzer S, Skrifvars MB, Castrén M. **Effect of an organizational change in a pre-hospital trauma care protocol and trauma transport directive in a large urban city: a before and after study.** Scand J Trauma Resusc Emerg Med 2 0 1 6 ; 9:26.

28. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. **A revision of the Trauma Score.** J Trauma 1989; 29:623-629.

29. Ocak G, Strums LM, Hoogeveen JM, Le Cessie S, Jukema GN. P**re-hospital identification of major trauma patients.** Langenbecks Arch Surg 2009; 394:285-292.

30. Vassar MJ, Holcroft JJ, Knudson MM, Kizer KW. **Fractures in access to and assessment of trauma systems.** J Am Coll Surg 2003; 197:717-725.

31. Chang DC, Bass RR, Cornwell EE, MacKenzie EJ. **Undertriage of elderly trauma patients to state-designated trauma centers.** Arch Surg 2008; 143:776-782.

32. Ichwan B, Darbha S, Shah MN et al. **Geriatric-specific triage criteria are more sensitive than standard adult criteria in identifying need for trauma center care in injured older adults.** Ann Emerg Med 2015;65:92-100.

33. Teasdale G, Maas A, Lecky F, Manley G, Stocchetti N, Murray G. **The Glasgow Coma Scale at 40 years: standing the test of time.** Lancet Neurol 2014; 13:844-854.

34. Hway Wong T, Krishnaswamy G, Vikas Nadkarni N et al. **Combining the new injury severity score with an anatomical polytrauma injury variable predicts mortality better than the new injury severity score and the injury severity score: a retrospective cohort study.** Scand J Trauma Resusc Emerg Med 2016;25.

35. Lefering R, Huber-Wagnet S, Nienaber U, Maegele M, Bouillon B. U**pdate of the trauma risk adjustment model of the TraumaRegister DGU™: the Revised Injury Severity Classification, version II.** Crit Care 2014; 18:476.

36. Scheetz LJ. **Comparison of type and severity of major injuries among undertriaged and correctly triaged older patients.** J Emerg Med 2012; 43:1020-1028.

37. Schoell SL, Doud AN, Weaver AA et al. **Development of a time sensitivity score for frequently occurring motor vehicle crash injuries.** J Am Coll Surg 2015; 220:305-312.

38. Cole TB. **Global road safety crisis remedy sought: 1.2 million killed, 50 million injured annually.** JAMA 2004; 2:2531-2532.

39. Leeflang MM, Bossuyt PM, Irwig L. **Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis.** J Clin Epidemiol 2009; 62:5-12.

40. Barnes J, Hassan A, Cuerden R, Cookson R, Kohlhofer J. **Comparison of injury severity between AIS 2005 and AIS 1990 in a large injury database.** Ann Adv Automot Med 2009; 53:83-89.

41. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2006.** Chicago, IL, 2006.

42. Hedges JR, Feero S, Moore B, Haver DW, Shultz B. **Comparison of pre-hospital trauma triage instruments in a semirural population.** J Emerg Med 1987; 5:197-208.

43. Newgard CD, Nelson MJ, Kampp MBS et al. **Out-of-hospital decision making and factors influencing the regional distribution of injured patients in a trauma system.** J Trauma 2011; 70:1345-1353.

44. Emerman CL, Shade B, Kubincanek J. **A comparison of EMT judgment and pre-hospital trauma triage instruments.** J Trauma 1991; 31:1369-1375.

45. Fries GR, McCalla G, Levitt MA, Cordova R. **A prospective comparison of paramedic judgment and the trauma triage rule in the pre-hospital setting.** Ann Emerg Med 1994; 24:885-889.

46. Newgard CD, Kampp MBS, Nelson MJ et al. **Deciphering the use and predictive value of "emergency medical services provider judgment" in out-of-hospital trauma triage: a multisite, mixed methods assessment.** J Trauma Acute Care Surg 2012; 72:1239-1248.

47. Harmsen AM, Giannakopoulos GF, Moerbeek PR, Jansma EP, Bonjer HJ, Bloemers FW. **The influence of pre-hospital time on trauma patients outcome: a systematic review.** Injury 2015; 46:602-609.

APPENDIX

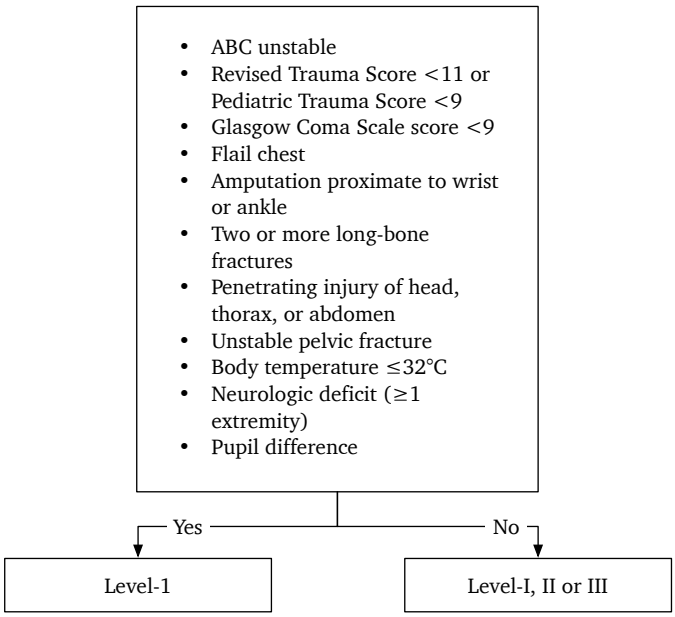


Figure 1 Abstraction of The National Protocol of Ambulance Services (Triage Choice of Hospital, version 7.1)

CHAPTER VII

Development and validation of a novel prediction model to identify patients in need of specialized trauma care during field triage: design and rationale of the GOAT study

Diagnostic & Prognostic Research

Rogier van der Sluijs; Thomas P.A. Debray;
Martijn Poeze; Loek P.H. Leenen; Mark van Heijl

ABSTRACT

Background

Adequate field triage of trauma patients is crucial to transport patients to the right hospital. Mistriage and subsequent inter-hospital transfers should be minimized to reduce avoidable mortality, life-long disabilities, and costs. Availability of a pre-hospital triage tool may help to identify patients in need of specialized trauma care and to determine the optimal transportation destination.

Methods

The GOAT (Gradient Boosted Trauma Triage) study is a prospective, multi-site, cross-sectional diagnostic study. Patients transported by at least five ground Emergency Medical Services to any receiving hospital within the Netherlands are eligible for inclusion. The reference standards for the need of specialized trauma care are an Injury Severity Score of 16 or greater and early critical-resource use, which will both be assessed by trauma data managers after the final diagnosis is made. Variable selection will be based on ease of use in practice and clinical expertise. A gradient boosting decision tree algorithm will be used to develop the prediction model. Model accuracy will be assessed in terms of discrimination (c-statistic) and calibration (intercept, slope, and plot) on individual participant data from each participating cluster (i.e., Emergency Medical Service) through internal-external cross-validation. A reference model will be externally validated on each cluster as well. The resulting model statistics will be investigated, compared, and summarized through an individual participant data meta-analysis.

Discussion

The GOAT study protocol describes the development of a new prediction model for identifying patients in need of specialized trauma care. The aim is to attain acceptable undertriage rates and to minimize mortality rates and life-long disabilities.

INTRODUCTION

Pre-hospital trauma triage is essential to get the right patient to the right hospital.[1] Erroneously transporting a patient requiring specialized trauma care to a lower-level trauma center is associated with higher mortality rates.[2,3] Conversely, transporting a patient not in need of specialized trauma care to a higher-level trauma center results in extra costs and overutilization of resources. These key metrics for triage quality are termed undertriage and overtriage, respectively. The American College of Surgeons Committee on Trauma guidelines state that trauma systems must aim to attain a maximum of 5% undertriage.[1]

One key component in the diagnostic strategy that determines the initial transportation destination is the use of a pre-hospital triage tool. These tools often involve the use of a prediction model or a flowchart where fulfillment of one of multiple criteria indicates the need for specialized trauma care. Unfortunately, a recent systematic review identified that the discriminative ability of many existing tools is quite poor.[4] One of the reasons is that simplification is key to facilitate their usefulness in clinical practice, thereby degrading predictive accuracy. There is limited time to collect patient data on-scene, and diagnostic modalities are very limited compared to hospitals.

The Trauma Triage App (TTApp) was recently developed to overcome the typical trade-off between simplicity and predictive accuracy. This mobile application implements a logistic regression model to estimate the need of specialized trauma care and provides an easy-to-use interface. This (reference) model was developed using individual participant data (IPD) from a single Emergency Medical Service (EMS) in the Netherlands. When the model was externally validated in a different EMS in the Netherlands, we found an undertriage rate of approximately 11%, at cost of approximately 50% overtriage.[5]

Although the reference model outperformed other tools, its discriminative value and generalizability could potentially be improved using a machine learning algorithm, a greater amount of IPD, participating EMSs and hospitals, and a more robust development strategy.[5]

In particular, the relatively small sample size (4950 patients, with 435 patients in need of specialized trauma care) limited the use of interaction terms and non-linear effects for modeling the included predictors and prevented any insight into the model's generalizability across different EMSs in the Netherlands. Therefore, the aims of the GOAT (Gradient Boosted Trauma Triage) study are (1) to develop a new prediction model on nationwide

IPD that accurately identifies patients in need of specialized trauma care in a pre-hospital setting, (2) to validate this prediction model on IPD from multiple EMSs during development, (3) to investigate sources of heterogeneity in model performance, and (4) to compare it to the reference model used in the initial version of the TTApp.

METHODS/DESIGN

Study Design

This is a prospective, multi-site, cross-sectional diagnostic study that is conducted to predict the need of specialized trauma care during field triage. We will adhere to existing recommendations on diagnostic model development, IPD meta-analysis (IPD-MA), and report the resulting model in accordance with the Transparent Reporting of a multivariable model for Individual Prognosis or Diagnosis (TRIPOD) guidelines.[6-9] Data collection started at Jan 1, 2015, and ended at Dec 31, 2018.

Participants

All patients, suspected of injury, transported by a ground EMS from the scene of injury to any emergency department in the Netherlands will be potentially eligible. The Netherlands is divided into 25 different EMS regions and 11 inclusive trauma regions. At least five different EMS regions will be included. These EMS regions have to be representative for urban, suburban, and rural areas. All hospitals, and consequently all trauma regions, with receiving emergency departments in the Netherlands collect the required patient outcomes and participate in this study.

Data Collection

Two distinct data sources will be merged to create a final dataset. These data sources consist of pre-hospital run reports, collected in a standardized manner by multiple EMSs, and the Dutch National Trauma Registry (in Dutch, *Landelijke Trauma Registratie* [LTR]). Run reports used by included EMSs are based on the template of the Basic Set of Ambulance Care (in Dutch, *Basisset Ambulancezorg* [BSA]) and include demographics, physiologic characteristics, mechanism of injury, injuries, patient status, on-scene treatments, initial transportation destination, and more. The LTR is a nationwide registry that collects patient data in accordance with an extended version of the Utstein registry template for uniform reporting of data following major trauma.[10] This registry covers all trauma-related hospital admissions of trauma-receiving emergency departments in the Netherlands since 2015.[11-13] Relevant patient outcomes included in this registry are, among others, Injury Severity Scores (ISS), early critical-resource use, intensive care unit admission, and death. Patient identification numbers used by EMSs are collected when available.

A combined deterministic and probabilistic linkage scheme will be used to match pre-hospital run reports and data from the LTR. Records are deterministically linked when pre-hospital patient identification numbers are available in both datasets. A probabilistic approach will be used to match patient records when unique identifiers are lacking. This approach utilizes machine learning methods and distance functions to identify matching records. Patients discharged directly from the emergency department are presumed not to have any of the investigated patient outcomes. This assumption combined with linking hospital and pre-hospital records had a sensitivity of 99.7% (95% CI, 99.0–99.9) and specificity of 100.0% (95% CI, 99.7–100.0) in data from a previous study.[14] The full data collection and record linkage strategy are depicted in Figure 1.

Outcome

The primary outcome is an ISS of 16 or greater coded by trained trauma data managers within 30 days after the emergency department admission. This reference standard is based on the Abbreviated Injury Scale version 2005, update 2008, and is recommended by the American College of Surgeons Committee on Trauma to evaluate triage quality.[1] Treating patients with an ISS of 16 or greater in higher-level trauma centers is associated with lower mortality rates.[2,15,16] The ISS is an anatomical score that is calculated after the final diagnosis is made. Since it is based on anatomic criteria, it is assumed to be identical to the patient status on-scene and is thus used as a diagnostic reference standard.

Because ISS is not perfectly correlated with resource utilization, we included a secondary, resource-based outcome measure to define the need for specialized trauma care.[17,18] The secondary outcome is early critical-resource use, which is a composite endpoint consisting of intubation in the pre-hospital setting, major surgical intervention, radiological intervention, or death within 24 h, as well as discharge to the intensive care unit from the emergency department. A similar endpoint is used in prior studies on pre-hospital trauma triage.[19]

Predictor Selection

Time is critical during field triage. Therefore, the number and complexity of hand-collected variables must be limited. To prevent the delay of definitive treatment, variables should be easily accessible during routine care, clearly defined, and measured in a standardized and reproducible way to improve transferability and predictive stability.[8] The candidate variables for model development were predefined based on prior evidence and clinical reasoning (Table 1). For instance, many candidate variables are criteria from the Field Triage Decision Scheme, which is the primary triage tool used by EMSs in the US.[1] The final set of variables will be selected prior to model development. The selection of variables is therefore independent of their performance in the training data. Additional

predictors (i.e., features), which are not predefined, will be engineered from these variables (e.g., the date of injury might be converted to three predictors indicating the day of the week, the current month, and the current season of the year).

The TTApp allows prediction models to use additional variables collected by the device on which the algorithm is embedded. These variables do not delay treatment since collection is computerized. Traveling times, global positioning systems locations, date, and time are variables that might provide extra predictive power to the hand-collected variables. Many predictors can be engineered from these variables, such as the season of the year, day of the week, regions, daytime or night, and more. No constraints are posed on the number and type of predictors that can be derived from these variables during the development phase.

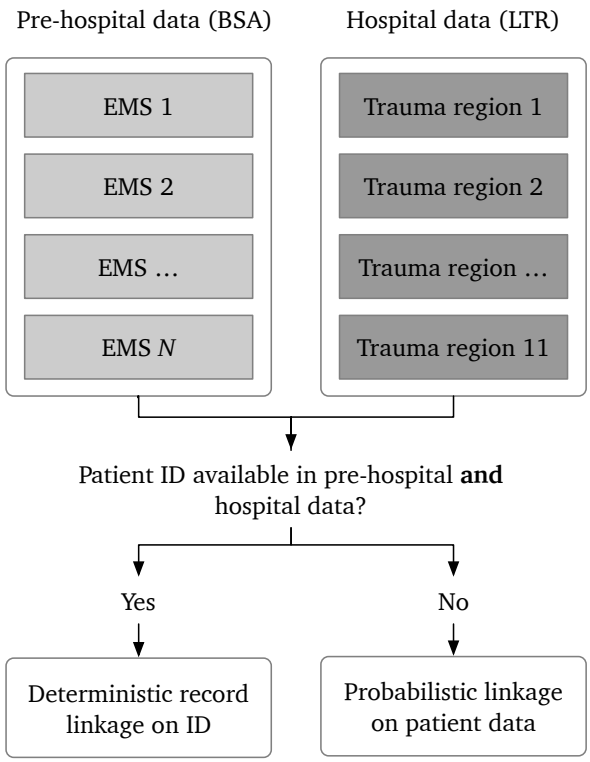


Figure 1 Data Collection and Record Linkage. Abbreviations: BSA, *Basic Set of Ambulance Care*; EMS, *Emergency Medical Service*; LTR, *Dutch National Trauma Registry*.

Missing Data

Most prediction modeling methods, such as logistic regression, are not able to deal with missing values and therefore require special care during development, validation, and implementation. For trauma triage, missing values are a particular concern because there may not always be time to measure critical variables. For this reason, we here adopt gradient boosting decision trees for prediction model development, as resulting prediction models can deal with missing values upon implementation. Briefly, decision tree algorithms implement surrogate splits for predictors with missing values and loosely operate under a missing-at-random assumption (as splits are conditional on some of the observed data). This yields an advantage in real-life situations, where pre-hospital data are often not fully available and surrogate splits can therefore be used to obtain an individual prediction nevertheless.

Multiple imputation will be used to address missing variables in the dataset in order to validate the reference model (which cannot accommodate for missing values). We will adopt multiple imputation methods that account for clustering across sites. Fifty different imputed datasets will be generated using chained equations by the R-package MICEMD.[20,21] Analyses will be applied to each individual dataset. Results will be averaged to provide point estimates. Confidence intervals will be calculated according to Rubin's rules.[22]

Statistical Analysis Methods

In this study, we will develop a gradient boosting decision tree with the LightGBM Python library, and we will compare it to the reference model by the means of internal-external cross-validation.[23-25]

Boosting is an ensemble technique that involves the estimation of multiple, related, prediction models.[26] The core concept of boosting is to add new models to the ensemble sequentially, in contrast to other ensemble strategies. Each model added to the ensemble is trained with respect to the error of the previously estimated models. Boosting can be applied to various families of prediction models and is often used in conjecture with decision trees.[27] The LightGBM Python library extends the boosting principle with various tunable hyperparameters (e.g., maximum tree depth, number of boosting iterations, custom objective functions) and regularization methods (e.g., subsampling a ratio of columns when constructing a new tree). Furthermore, it deals with missing data by sparsity-aware split finding. The default direction of a node is learned in the tree construction process, so that it minimizes the error in the training data.

Table 1 Candidate Variables for Predictor Engineering

Variable	Reason for inclusion
Demographics	
Age	Included in the FTDS
Gender	Associated with the reference standard in previous research and interacts with other candidate variables
Vital signs	
Glasgow Coma Scale, eyes component	Included in the FTDS
Glasgow Coma Scale, motor component	Included in the FTDS
Glasgow Coma Scale, verbal component	Included in the FTDS
Systolic blood pressure	Included in the FTDS
Diastolic blood pressure	Expected interactions with other candidate variables (e.g., systolic blood pressure)
Heart rate	Expected interactions with other candidate variables (e.g., systolic blood pressure)
Respiratory rate	Included in the FTDS
Intubation	Direct indication of resource-use
Oxygen saturation	Associated with the reference standard in previous research and expected interactions with other candidate variables
Mechanism of injury	
MVA, (excl. motorcycles, mopeds, scooters)	Included in the FTDS
Motorcycle accident	Included in the FTDS
Moped, scooter accident	
MVA, pedestrian	Included in the FTDS
MVA, different	Included in the FTDS
Gunshot	Expected association with the reference standard and other candidate variables (e.g., penetrating injury)
Stab wound	Expected association with the reference standard and other candidate variables (e.g., penetrating injury)
Struck with blunt object	Expected association with the reference standard
Fall, same level	Included in the FTDS
Fall, higher level	Included in the FTDS
Asphyxia	Associated with the reference standard in previous research
Burns, % of body surface	Associated with the reference standard in previous research

Variable	Reason for inclusion
Injury type	
Penetrating injury to head, neck, torso, and extremities proximal to elbow and knee	Included in the FTDS
Flail chest	Included in the FTDS
Paralysis	Included in the FTDS
Open or depressed skull fracture	Included in the FTDS

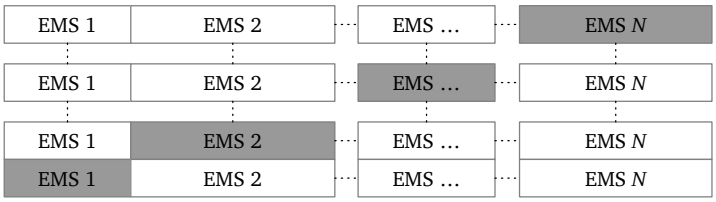
Abbreviations: FTDS, *Field Triage Decision Scheme*; MVA, *motor vehicle accident*.

A robust model development strategy will be implemented to avoid model optimism. First, internal-external cross-validation (IECV) will be used generate N pairs of development and (non-random) validation samples, where N is the number of participating clusters (EMSs). This technique iteratively uses IPD from N–1 clusters to develop a prediction model and the remaining cluster's IPD for its external validation. This yields N scenarios in which model performance can be investigated in an independent sample and compared to the reference model. A major difference with traditional cross-validation is that hold-out samples in IECV are non-random if the available clusters differ from one another, which allows to assess model generalizability (rather than reproducibility).

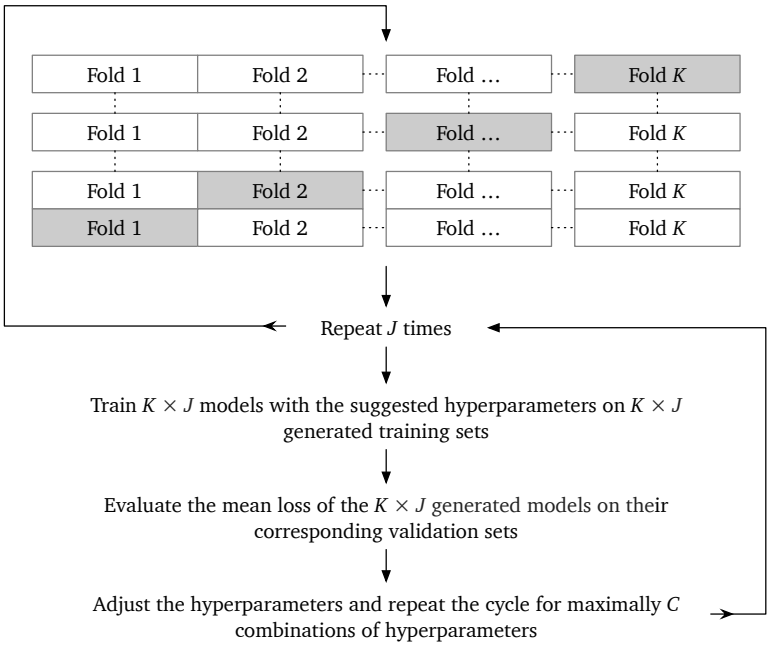
In each of the N training datasets, we will develop a prediction model using LightGBM. The set of predefined hyperparameters will be optimized for each model using ten iterations of stratified tenfold cross-validation with a shuffle prior to each iteration (see Table 2). Hereto, we will adopt a Tree-structured Parzen Estimator algorithm to minimize the mean squared error within a restricted search space in 500 iterations.[28] We limited the amount of hyperparameters to be optimized to avoid overfitting and to enable more extensive modeling of individual predictors.

Second, in each IECV round, we will externally validate the developed model in the test sample and assess its discrimination (c-statistic) and calibration (intercept, slope, and plot) performance. Two scenarios will be explored, one including class weights that are inversely proportional to the outcome occurrence in the development data and one without class rebalancing. We will also assess its comparative performance with the reference model, by quantifying the difference in c-statistic and performing decision curve analysis.[29]

Step 1 - Internal-external cross-validation: create *N* training and test sets, where *N* is the number of clusters (EMSs).



Step 2 - Hyperparameter optimization: for every training set generated in step 1, define the optimal hyperparameters by iterated stratified *K*-fold cross-validation with shuffling prior to each iteration:



Step 3 - Model development: train one model for every training set generated in step 1 with the optimal hyperparameters found by step 2.

Step 4 - Model evaluation: evaluate the *N* resulting models by assessing their discrimination (e.g., c-statistic) and calibration (e.g., intercept, slope) on their paired test set generated in step 1.

Step 5 - Final model: repeat step 2 and 3 on the full dataset to create a final model (i.e., find the optimal hyperparameters on the full dataset and train a final model with these parameters on the full dataset).

Figure 2 Model Development Strategy. Abbreviations: EMS, *Emergency Medical Service*.

Table 2 Hyperparameters

Parameter	Explanation
Free	
Lambda L1	Shrinkage rate (how much will the weights be adjusted every iteration)
Number of leaves	Maximum number of leaves in one tree
Lambda L1	L1 regularization
Lambda L2	L2 regularization
Feature fraction	Randomly select part of the predictors on each iteration
Fixed	
Early stopping	The cross-validation score needs to improve at least every *n* rounds to continue with the next boosting iteration
Maximum depth	Maximum tree depth (note that its less relevant here since the tree grows leaf-wise)
Minimum data	Minimal number of records in one leaf. A higher number prevents overfitting
Bagging fraction	Randomly select part of the data without resampling
Bagging frequency	Per how many rounds should bagging be applied
Unbalanced data	Does data need to be balanced or not

The third and final step will be to construct one model based on the complete dataset. The full model development strategy is illustrated in Figure 2. Estimates of model discrimination (c-statistic) and calibration (intercept and slope) from all hold-out samples (i.e., the different clusters) will be pooled separately by IPD-MA for both the reference model and the newly developed model. Random effects meta-analysis models, in which the weights are based on the within and between-cluster error variance, will be used to account for heterogeneity between the available clusters.[30] The between-study standard deviation will be reported from the IPD-MA. Restricted maximum likelihood estimation will be applied to estimate variance components, and the Hartung-Knapp-Sidik-Jonkman method will be used to derive 95% confidence intervals for the summary estimates of model performance.[30]

DISCUSSION

Trauma systems can only reach their full potential when patients are transported to the right hospitals within the right time. Mortality rates, morbidity rates, and costs can be potentially reduced by minimizing undertriage, overtriage, and inter-hospital transfer rates. A pre-hospital triage tool is crucial to aid EMS professionals in order to achieve this goal.

The TTApp provides a digital platform that is easy to use, fast, and capable of incorporating complex prediction models, and provides the possibility for iterative improvements. The new prediction model proposed in this study protocol aims to improve predictive accuracy and generalizability through a robust model development strategy.

Limitations

One key element of trauma systems is centralization, which should enable the most efficient use of finite resources. Centralization and its positive consequences (i.e., high-volume trauma centers) are known to lower mortality rates. One limitation of the primary outcome is the use of an ISS of 16 or greater as the reference standard for the need of specialized trauma care, since the ISS is a scale that does not perfectly correlate with resource use.[17,18] The secondary outcome eliminates this limitation, but is not officially used to evaluate triage accuracy.[1]

A second limitation is that we focus on gradient boosting decision tree and do not evaluate other prediction modeling strategies. However, we do not aim to develop a perfect prediction model (which is impossible anyhow) and believe that the size of our dataset, the restriction of unknown hyperparameters, and the implementation of regularization will prevent overfitting. Furthermore, by avoiding additional comparisons with other modeling strategies, we effectively minimize the danger of chance findings and overoptimism. Finally, it is important to realize that we chose to avoid regression analysis as the current prediction model for trauma triage (which is based on logistic regression) suffers from missing values in clinical practice, a problem that is remedied by adopting gradient boosting models.

A third limitation of this study is the use of frequentist meta-analysis methods to evaluate model performance in new settings and populations. In this regard, the estimation of between-cluster heterogeneity and prediction intervals may benefit from adopting a Bayesian approach.[31]

Implications

The TTApp is currently implemented at multiple EMSs in the Netherlands. This existing infrastructure allows us to replace the reference model with the newly developed model if it proves to be better. A software update will then implement the new prediction model on the currently used devices, so that the new model can be used almost instantly. Higher predictive accuracy and better generalizability of the TTApp will likely lead to reduced mistriage rates and, as a consequence, lower mortality rates and less life-long disabilities. The final model will be made available as a Python object through the supplementary content.

Conclusions

The TTApp is currently used by multiple EMSs in the Netherlands to provide EMS professionals with decision support during field triage. This study protocol outlines the methodology that will be used to construct an improved prediction model, with emphasis on high predictive accuracy and broad generalizability.

Acknowledgements

The authors would like to thank L.M. Sturms from the National Network of Acute Care (in Dutch, *Landelijke Netwerk Acute Zorg*) for her contribution to the data collection and W.M.J. Hoogeveen and G. van Duin from Ambulance Care Netherlands (in Dutch, *Ambulancezorg Nederland*) for their contributions to the study protocol.

Author Contributions

RS contributed to the study design and statistical analysis, and drafted and revised the manuscript. TPAD contributed to the study design and statistical analysis plan and revised the draft paper. MP, MH, and LPHL contributed to the study design and read, revised, and approved the final manuscript.

REFERENCES

1. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Jul 27, 2017).

2. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

3. Staudenmayer K, Weiser TG, Maggio PM, Spain DA, Hsia RY. **Trauma center care is associated with reduced readmissions after injury.** J Trauma Acute Care Surg 2016; 80(3): 412-6; discussion 6-8.

4. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld RA, Leenen LPH, van Heijl M. **Accuracy of pre-hospital triage protocols in selecting severely injured patients: A systematic review.** J Trauma Acute Care Surg 2017; 83(2): 328-39.

5. van Rein EAJ, van der Sluijs R, Voskens FJ, et al. **Development and validation of a prediction model for pre-hospital triage of trauma patients.** JAMA Surg 2019.

6. Collins GS, Reitsma JB, Altman DG, Moons KG. **Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement.** BMJ 2015; 350: g7594.

7. Moons KG, Kengne AP, Grobbee DE, et al. **Risk prediction models: II. External validation, model updating, and impact assessment.** Heart 2012; 98(9): 691-8.

8. Moons KG, Kengne AP, Woodward M, et al. **Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker.** Heart 2012; 98(9): 683-90.

9. Debray TP, Moons KG, Ahmed I, Koffijberg H, Riley RD. **A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis.** Stat Med 2013; 32(18): 3158-80.

10. Dick WF, Baskett PJ, Grande C, et al. **Recommendations for uniform reporting of data following major trauma--the Utstein Style.** An International Trauma Anaesthesia and Critical Care Society (ITACCS) initiative. Eur J Emerg Med 1999; 6(4): 369-87.

11. Champion HR, Copes WS, Sacco WJ, et al. **The Major Trauma Outcome Study: establishing national norms for trauma care.** J Trauma 1990; 30(11): 1356-65.

12. Landelijk Netwerk Acute Zorg. **Landelijke Traumaregistratie 2011-2015: Rapportage Nederland.** Utrecht, 2016.

13. Ringdal KG, Coats TJ, Lefering R, et al. **The Utstein template for uniform reporting of data following major trauma: a joint revision by SCANTEM, TARN, DGU-TR and RITG.** Scand J Trauma Resusc Emerg Med 2008; 16: 7.

14. Voskens FJ, van Rein EAJ, van der Sluijs R, et al. **Accuracy of pre-hospital triage in selecting severely injured trauma patients.** JAMA Surg 2018; 153(4): 322-7.

15. 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(1): 173-8.

16. Nathens AB, Jurkovich GJ, Maier RV, et al. **Relationship between trauma center volume and outcomes.** JAMA 2001; 285(9): 1164-71.

17. 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.

18. Newgard CD, Hedges JR, Diggs B, Mullins RJ. **Establishing the need for trauma center care: anatomic injury or resource use?** Prehosp Emerg Care 2008; 12(4): 451-8.

19. Newgard CD, Fu R, Lerner EB, et al. **Deaths and high-risk trauma patients missed by standard trauma data sources.** J Trauma Acute Care Surg 2017; 83(3): 427-37.

20. Van Buuren S, Groothuis-Oudshoorn K. **MICE: Multivariate Imputation by Chained Equations in R.** Journal of Statistical Software 2011; 45(3): 1-67.

21. Audigier V, White I, Jolani S, et al. **Multiple imputation for multilevel data with continuous and binary variables.** Stat Methods Med Res 2018.

22. Marshall A, Altman DG, Holder RL, Royston P. **Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines.** BMC Med Res Methodol 2009; 9: 57.

23. Moons KG, Donders AR, Steyerberg EW, Harrell FE. **Penalized maximum likelihood estimation to directly adjust diagnostic and prognostic prediction models for overoptimism: a clinical example.** J Clin Epidemiol 2004; 57(12): 1262-70.

24. Chen T, Guestrin C. **XGBoost: A scalable tree boosting system.** arXiv:160302754 2016.

25. Ke G, Meng Q, Finley T, et al. **LightGBM: A Highly Efficient Gradient Boosting Decision Tree.** Conference on Neural Information Processing Systems (NIPS). Long Beach, CA, USA; 2017.

26. Friedman JH. **Greedy function approximation: A gradient boosting machine.** Ann Statist 2001; 29(5): 1189-232.

27. Natekin A, Knoll A. **Gradient boosting machines, a tutorial.** Front Neurorobot 2013; 7: 21.

28. Bergstra J, Bardenet R, Bengio Y, Kégl B, R. Bardenet YBaBKgAfH-pOPNIPSN, 2546–2554, 2011. **Algorithms for Hyper-parameter Optimization.** Proc Neural Information Processing Systems 2011; 24: 2546-54.

29. Vickers AJ, Elkin EB. **Decision curve analysis: a novel method for evaluating prediction models.** Med Decis Making 2006; 26(6): 565-74.

30. Debray TP, Damen JA, Snell KI, et al. **A guide to systematic review and meta-analysis of prediction model performance.** BMJ 2017; 356: i6460.

31. Sutton AJ, Abrams KR. **Bayesian methods in meta-analysis and evidence synthesis.** Stat Methods Med Res 2001; 10(4): 277-303.

CHAPTER 8

Development and validation of pre-hospital prediction models to select patients in need of specialized trauma care (GOAT study): an individual participant data meta-analysis

Rogier van der Sluijs; Thomas P.A. Debray; Job F. Waalwijk;
Robin D. Lokerman; Martijn Poeze; Luke P.H. Leenen; Mark van Heijl;
for the Pre-hospital Trauma Triage Research Collaborative

ABSTRACT

Objectives

To develop and validate prediction models to identify patients in need of specialized trauma care in the pre-hospital setting using simple, routinely available, and automatically measured risk factors.

Design

Individual participant data meta-analysis.

Setting

Populations from six Emergency Medical Services, 55 hospitals, and seven inclusive trauma regions, the Netherlands.

Participants

A total of 133,196 of patients with suspected injuries during field triage were eligible for participation. Included patients were transported by ambulance with high priority from the scene of injury to a trauma-receiving emergency department between Jan 2015 and Dec 2017.

Main Outcome Measures

Two gradient-boosting decision trees were developed to predict (i) injury severity and (ii) in-hospital resource use during field triage based on individual participant data. Particular emphasis was placed on the use of predictors that are easily obtained on site, and on the capability to deal with missing values. Internal-external cross-validation and random-effect meta-analyses were performed to examine optimism and generalizability of model predictions. Decision curve analysis was used to compare model performance and net benefit with contemporary triage tools in a random sub-sample.

Results

Model construction and validation was based on a prospective cohort of consecutive patients from six regions. Both models achieved an excellent discrimination and calibration performance. The model to assess injury severity yielded a pooled summary c-statistic of 0.84, a pooled O:E ratio of 1.00, and a pooled calibration slope of 1.01. The second model to predict resource use yielded a c-statistic of 0.88, an O:E ratio of 1.01, and a calibration slope of 0.99. The c-statistic for both models was 0.07 (14% absolute improvement) higher than the c-statistic of the second-best triage tool. Decision curve-analyses demonstrated superior net benefit for risk thresholds below 10%, and similar net benefit for higher thresholds.

Conclusions
The developed models were highly accurate, superior to triage tools used in contemporary practice, and were specifically designed to generate predictions in a time-critical setting with incomplete data availability. Well-calibrated predictions now enable a precision-medicine approach to get the right patient to the right hospital.

INTRODUCTION

Adequate field triage is paramount to centralize severely injured patients within inclusive trauma systems. Mistriage can be detrimental in trauma systems with a high degree of centralization of resources and expertise. Undertriage – transporting patients in need of specialized trauma care to lower-level trauma centers – is associated with increased mortality rates, whereas overtriage – transporting mildly injured patients to higher-level trauma centers – induces resource scarcity and is not considered to be cost-effective.[1,2] Inclusive trauma systems should aim to achieve undertriage rates of <5% according to the guidelines of the American College of Surgeons Committee on Trauma.[3] A systematic review on worldwide triage accuracy, however, indicated that no inclusive trauma system was able to comply with this guideline while maintaining acceptable overtriage rates.[3,4]

One key component to get the right patient to the right hospital is the pre-hospital triage protocol. Simple flowcharts such as the Field Triage Decision Scheme of the American College of Surgeons Committee on Trauma and the Dutch National Protocol of Ambulance Services are extensively used, but have limited predictive accuracy.[5,6] On the other hand, more complex models (e.g., the Trauma Score) fall into disuse in time-critical situations.[7] Resolutions to this typical trade-off are computerized triage tools such as the recently developed Trauma Triage App.[8] Computerized systems enable Emergency Medical Services (EMS) professionals to use complex predictions models through user-friendly interfaces. However, existing tools still require substantial human input and complete availability of input variables. In addition, their generalizability across different trauma populations remains unclear due to the scarcity of validation studies.

Model performance and usefulness in daily practice could potentially be improved by combining large amounts of individual participant data (IPD) from multiple EMS regions, by adopting machine learning algorithms for model development, and by assessing the generalizability of model predictions across multiple settings and locations. The GOAT (Gradient Boosted Trauma Triage) study was designed to achieve these goals.[9] In particular, the aims of this study were (1) to develop and validate a prediction model on nationwide IPD to identify patients in need of specialized trauma care during field triage,

(2) to facilitate dynamic risk predictions based on incomplete data, (3) to investigate sources of heterogeneity in model performance, and (4) to compare it to usual care in contemporary practice.

METHODS

Source of Data

The GOAT study was a prospective, multi-region, diagnostic study to develop and validate prediction models aimed at selecting patients in need of specialized trauma care. A protocol describing the design and methodology of this study was previously published.[9] We planned to develop and validate two prediction models on IPD from six regions of the Trauma Triage Continuum of Care Cohort, that captured and combined data from a consecutive series of pre-hospital electronic health records linked to the Dutch National Trauma Registry from Jan 1, 2015 until Dec 31, 2017.[9] The study was reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis guidelines and specific recommendations for IPD meta-analyses.[10,11]

Participants

All patients, suspected of injury, transported by six EMSs (Amsterdam-Amsteland, Brabant Midden-West, Brabant Noord, Gelderland-Zuid, Rotterdam-Rijnmond, and Utrecht) with high priority from the scene of injury to any emergency department within seven out of 11 inclusive trauma regions in the Netherlands were eligible for inclusion.

Approximately 500,000 patients are transported by these EMSs annually, within a region that covers almost 7000 km^2 with a population of over 5.7 million people. These regions encompass six higher-level trauma centers and 25 lower-level trauma centers. All higher-level trauma centers feature an intensive care unit (ICU) and offer trauma care at the highest level available within the trauma system. EMS regions were considered to be the cluster-level in this study, as EMSs hold the direct responsibility for ensuring adequate field triage.

Outcome

The primary outcome was an Injury Severity Score (ISS) of 16 or greater as coded by trained trauma data managers after a patient's final diagnosis was made.[12] The ISS was calculated based on the Abbreviated Injury Scale version 2005, update 2008.[13] The American College of Surgeons Committee on Trauma suggests using this reference standard in order to evaluate inclusive trauma systems.[3]

The secondary outcome was early critical-resource use, which was a composite endpoint consisting of pre-hospital intubation, major surgical or radiological intervention, discharge to the ICU from the emergency department, or death within 24 h after hospital admission (see Appendix). Resource-based reference standards directly relate to one fundamental aspect of inclusive trauma systems (i.e., centralization of resources) and are thus frequently proposed as an alternative to evaluate triage accuracy.[5]

Predictor Selection

Time is extremely limited in an acute care setting. This study restricted candidate variables for model development to those that were automatically collected by computerized devices, assessed by the dispatch center (i.e., variables that were available before arrival at the scene of injury), or routinely collected by EMS professionals. The final selection of variables was based on the aforementioned criteria, existing triage tools, and clinical reasoning (Table 1).

Table 1 Variables Used for Predictor Engineering

Variable	Measurement
General	
Age	Dispatch center or EMS professionals
Gender	Dispatch center or EMS professionals
Date/time	Computerized
Medical doctor on-scene	Dispatch center
Ambulance response time	Computerized
Ambulance priority	Dispatch center
Vital signs	
Glasgow Coma Scale score	EMS professionals
Blood pressure	Computerized
Heart rate	Computerized
Respiratory rate	EMS professionals
Oxygen saturation	Computerized
Heart rate	Computerized
Intubation	Computerized

Abbreviations: EMS, *Emergency Medical Service*.

Seventeen predictors were engineered from the selected variables. Dates were transformed to weekdays, time was converted to hour of the day, and vital signs were all summarized to their respective first measurements.

Missing Data

Missing data are empirical in an acute care setting, thereby hindering the use of models that rely on complete data availability, such as logistic regression. The Gradient Boosting Decision Tree (GBDT) algorithm, used to construct the models in this investigation, handles missing data by design.[14] It does not necessitate case-wise deletion, nor does it require missing values to be (multiply) imputed upon development and implementation. Instead, when a predictor value is missing, the GDBT algorithm determines the split-direction of a node that yields the minimal training loss in a process termed sparsity-aware split finding.[14]

We did adopt a multi-level multiple imputation strategy to compute baseline characteristics and to enable comparative analysis between the newly developed models and triage tools implemented in usual care. Forty-eight different imputed datasets were generated using the R-package micemd (12 datasets per processor core).[15] Multi-level multiple imputation was used to account for cluster differences.

Statistical Analysis Methods

We aimed to develop and externally validate two GBDTs to select patients in need of specialized trauma care: one for the primary and one for the secondary outcome. These models were developed using the same strategy (including variable engineering and selection, estimation of decision trees, and hyperparameters).

First, internal-external cross-validation was used to generate six pairs of training and validation datasets.[16] The IPD of each participating EMS served as the validation dataset once, at which occasion the IPD from the other EMSs were used as the training data. Second, hyperparameters of the GBDTs – including the potential application of a positive scale factor to adjust for class imbalance – were optimized using two iterations of five-fold cross-validation for each training dataset. One model was created per training set using its unique set of optimized hyperparameters. Samples in the training sets were weighted to balance the effect of each of the EMSs' IPD on the final model during model construction. This yielded six scenarios in which we could investigate model performance in an independent (non-random) hold-out sample. Third, we performed random effects meta-analyses to assess overall model performance and inter-region heterogeneity. Finally, we repeated hyperparameter optimization on the complete dataset and constructed one final model for each outcome based on all available IPD.

The ability to differentiate between patients that did need specialized trauma care and those who did not (i.e., model discrimination) was assessed by means of the concordance (c)-statistic. Model calibration was assessed using the calibration intercepts, observed versus expected (O:E) ratios, and calibration slopes.

Model performance was compared to two triage tools used in contemporary practice: the National Protocol of Ambulance Services (Triage Choice of Hospital, version 8.1) and a prediction model that was incorporated into the Trauma Triage App.[8] A random sample of 3629 patients from three regions (Amsterdam-Amstelland, Rotterdam-Rijnmond, and Utrecht) was selected to perform decision curve analyses and to assess the external validity of the Trauma Triage App.[8] The traditional triage tools were compared to restricted versions of the models developed in this study. These restricted models were constructed using the same hyperparameters as the final models, but were exclusively based on IPD from the three remaining regions (Brabant Midden-West, Brabant Noord, and Gelderland Zuid). Additional variables of the Trauma Triage App that were not used in our new models (e.g., suspected injury types) were obtained from unstructured text fields similar to its derivation setting.[8] We quantified the difference in c-statistic between the Trauma Triage App and the final models on the same subset of IPD. Performance of the National Protocol of Ambulance Services was simulated using the accuracy metrics reported by Voskens and colleagues (see Appendix).[6]

Rubin's rules were applied to calculate 95% confidence intervals for multiply imputed estimates.[17] Confidence intervals for ratios and c-statistics were calculated using bootstrapped percentile intervals based on 2000 replications. All analyses were conducted in Python (v3.7) and R (v3.6.1), mainly using the packages [xgboost, scikit-learn, micemd, and metamisc]. Corresponding source code is available from the appendix.

Patient and Public Involvement

Patients were not directly involved in conducting this study. No patients were asked for advice on interpretation of the results or to revise the manuscript. We do not plan to involve patients in dissemination of the results at this moment.

RESULTS

Approximately 1.4 million electronic health records were identified between Jan 2015 and Dec 2017. Half of these patients was urgently transported to an emergency department of which 133,196 were suspected of injuries (see Appendix).

Table 2 Characteristics of the Included Patients

Variable	All patients (n=133,196)	EMS region					
		Rotterdam-Rijnmond (n=30,972)	Amsterdam-Amstelland (n=31,058)	Brabant Midden-West (n=27,171)	Brabant Noord (n=14,163)	Utrecht (n=19,123)	Gelderland Zuid (n=10,709)
General							
Age, y	61.5 (39.2 - 79.5)	61.7 (39.2 - 80.0)	56.4 (33.9 - 76.8)	63.0 (41.7 - 79.9)	65.1 (44.8 - 80.5)	62.9 (40.4 - 80.3)	64.4 (43.4 - 79.8)
Men	66,480 (49.9)	15,499 (50.0)	15,948 (51.3)	13,714 (50.5)	7038 (49.7)	9019 (47.2)	5262 (49.1)
Women	66,716 (50.1)	15,473 (50.0)	15,110 (48.7)	13,457 (49.5)	7125 (50.3)	10,104 (52.8)	5447 (50.9)
Medical doctor on-scene	3273 (2.5)	936 (3.0)	907 (2.9)	636 (2.3)	377 (2.7)	145 (0.8)	272 (2.5)
Ambulance response time	8.9 (6.0 - 12.4)	8.0 (5.5 - 11.4)	7.7 (5.5 - 11.0)	9.4 (6.7 - 12.8)	9.7 (6.8 – 13.0)	10.0 (7.0 - 14.0)	9.9 (6.8 - 13.6)
Highest priority	67,933 (51.0)	17,164 (55.4)	22,504 (72.5)	11,517 (42.4)	6156 (43.5)	6690 (35.0)	3901 (36.4)
Vital signs							
Systolic blood pressure, mm Hg	140 (125 - 159)	140 (125 - 160)	138 (122 - 155)	140 (125 - 158)	140 (125 - 159)	142 (127 - 161)	140 (125 - 160)
Systolic blood pressure <90 mm Hg	1857 (1.4)	401 (1.3)	495 (1.6)	336 (1.2)	198 (1.4)	276 (1.4)	151 (1.4)
Respiratory rate	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (14 - 18)	16 (13 - 18)	16 (14 - 18)	16 (14 - 18)
Respiratory rate <10 or >29 breaths/min	3064 (2.3)	766 (2.5)	819 (2.6)	501 (1.8)	296 (2.1)	395 (2.1)	287 (2.7)
Heart rate, bpm	81 (73 - 93)	83 (74 - 94)	83 (74 - 94)	81 (72 - 92)	80 (71 - 92)	80 (72 - 90)	80 (72 - 90)
Oxygen saturation	97 (96 - 99)	98 (96 - 99)	98 (96 - 99)	97 (96 - 98)	97 (95 - 98)	97 (95 - 99)	98 (96 - 99)
Glasgow Coma Scale score <14	7753 (5.8)	1767 (5.7)	1771 (5.7)	1415 (5.2)	757 (5.3)	1411 (7.4)	632 (5.9)

Outcomes							
In-hospital stay	39,766 (29.9)	8038 (26.0)	7272 (23.4)	8821 (32.5)	4728 (33.4)	6679 (34.9)	4228 (39.5)
Length of hospitalization*	4.0 (1.2 - 8.0)	3.8 (1.1 - 7.6)	3.0 (1.0 - 8.0)	4.0 (1.2 - 8.1)	4.0 (1.3 - 8.1)	4.7 (1.7 - 8.7)	4.0 (1.0 – 8.0)
Injury Severity Score*	9 (4 - 9)	9 (4 - 9)	6 (4 - 9)	8 (4 - 9)	9 (4 - 9)	9 (4 - 9)	6 (4 - 9)
Injury Severity Score ≥16*	3352 (8.4)	706 (8.7)	576 (7.9)	722 (8.1)	359 (7.6)	373 (8.8)	616 (9.2)
Injury Severity Score ≥3, per region*	19,480 (49.0)	3974 (49.4)	3335 (45.9)	4140 (46.9)	2371 (50.1)	3719 (55.7)	1941 (45.9)
Head and neck	3750 (9.4)	825 (10.3)	791 (10.9)	725 (8.2)	392 (8.3)	678 (10.2)	339 (8.0)
Face	69 (0.2)	23 (0.3)	4 (0.1)	15 (0.2)	7 (0.1)	15 (0.2)	5 (0.1)
Thorax	2653 (6.7)	555 (6.9)	435 (6.0)	502 (5.7)	317 (6.7)	560 (8.4)	284 (6.7)
Abdomen	547 (1.4)	136 (1.7)	87 (1.2)	100 (1.1)	59 (1.2)	93 (1.4)	72 (1.7)
Extremities	13,588 (34.2)	2668 (33.2)	2189 (30.1)	3023 (34.3)	1749 (37.0)	2590 (38.8)	1369 (32.4)
External	258 (0.6)	72 (0.9)	43 (0.6)	47 (0.5)	13 (0.3)	57 (0.9)	26 (0.6)
Early critical-resource use*	3470 (8.7)	770 (9.5)	582 (8.0)	717 (8.1)	342 (7.2)	674 (10.1)	385 (9.1)
Discharge from emergency department to ICU	2509 (6.3)	482 (6.0)	392 (5.4)	570 (6.4)	267 (5.6)	508 (7.6)	290 (6.9)
Out-of-hospital intubation	915 (2.3)	219 (2.7)	176 (2.4)	134 (1.5)	102 (2.2)	177 (2.7)	107 (2.5)
Major intervention <12 h	787 (2.0)	277 (3.4)	146 (2.0)	130 (1.5)	53 (1.1)	69 (1.6)	112 (1.7)
Mortality <24 h	380 (1.0)	68 (0.8)	80 (1.1)	91 (1.0)	31 (0.7)	43 (1.0)	67 (1.0)

Abbreviations: EMS, *Emergency Medical Service*; ICU, *intensive care unit*. Variables with missing values were multiply imputed. Data are represented as median, IQR or n (%). *Hospitalized patients only.

Baseline characteristics of the included patients are shown in Table 2. The median age in this cohort was 61.5 years (IQR, 39.2 – 79.5). The percentage of men and women was approximately even. Ambulance response times were similar between the six participating regions with a median of 8.9 (IQR, 6.0 – 12.4) minutes. The percentage of ambulances activated with highest priority compared to the total number of urgent ambulance activations ranged from 35% to 72% between regions.

The median systolic blood pressure in the cohort was 140 and 1857 (1%) patients had a systolic blood pressure lower than 90. Respiratory rates, heart rates, and oxygen saturation rates were similar between regions with medians of 16/min, 81/min, and 97%, respectively. Nearly eight thousand patients (6%) had an impaired Glasgow Coma Scale of less than 14.

Approximately 40,000 patients (30%) were hospitalized. The median length of stay for these patients was 4.0 days (IQR, 1.2 – 8.0). Median ISSs of hospitalized patients ranged from six to nine between the participating EMSs. A total of 3352 (8%) hospitalized patients was considered to be severely injured according to the primary outcome (ISS ≥16), whereas 3470 (9%) patients made use of critical-resources according to the secondary outcome.

Data were partially missing for systolic blood pressure (27%), respiratory rate (36%), heart rate (18%), oxygen saturation (29%), Glasgow Coma Scale (14%), response time (6%), and ambulance activation priority (4%). No data were systematically missing.

Model Performance

Results from the univariate random effects meta-analyses are depicted in Figure 1 and Table 3. The pooled estimates of the c-statistics were 0.84 (95%-CI, 0.82 – 0.87) for the primary outcome model, and 0.88 (95% CI, 0.86 – 0.90) for the secondary outcome model. The pooled summary estimates of the total O:E ratio were 1.00 (95%-CI, 0.69 – 1.30) and 1.01 (95%-CI, 0.74 – 1.28), respectively. Although the 95% prediction intervals (95% PI) for the total O:E ratio ranged from 0.13 to 1.86 and 0.25 to 1.77, the actual estimates of between-region heterogeneity were relatively low (tau for log OE ratio = 0.08 and 0.06). Calibration slopes were homogeneous between EMSs (Figure 1) with pooled estimates of 1.01 (95%-CI, 0.98 – 1.05) and 0.99 (95%-CI, 0.96 – 1.03) for the primary and, respectively, secondary outcome models. Additional results are presented in the appendix, and include calibration plots and the O:E meta-analysis.

Because we observed potential heterogeneity in the overall calibration (as reflected by the prediction intervals of the total O:E ratio), we recalibrated the final models' predictions using a simple logistic regression model where the linear predictor was included as an

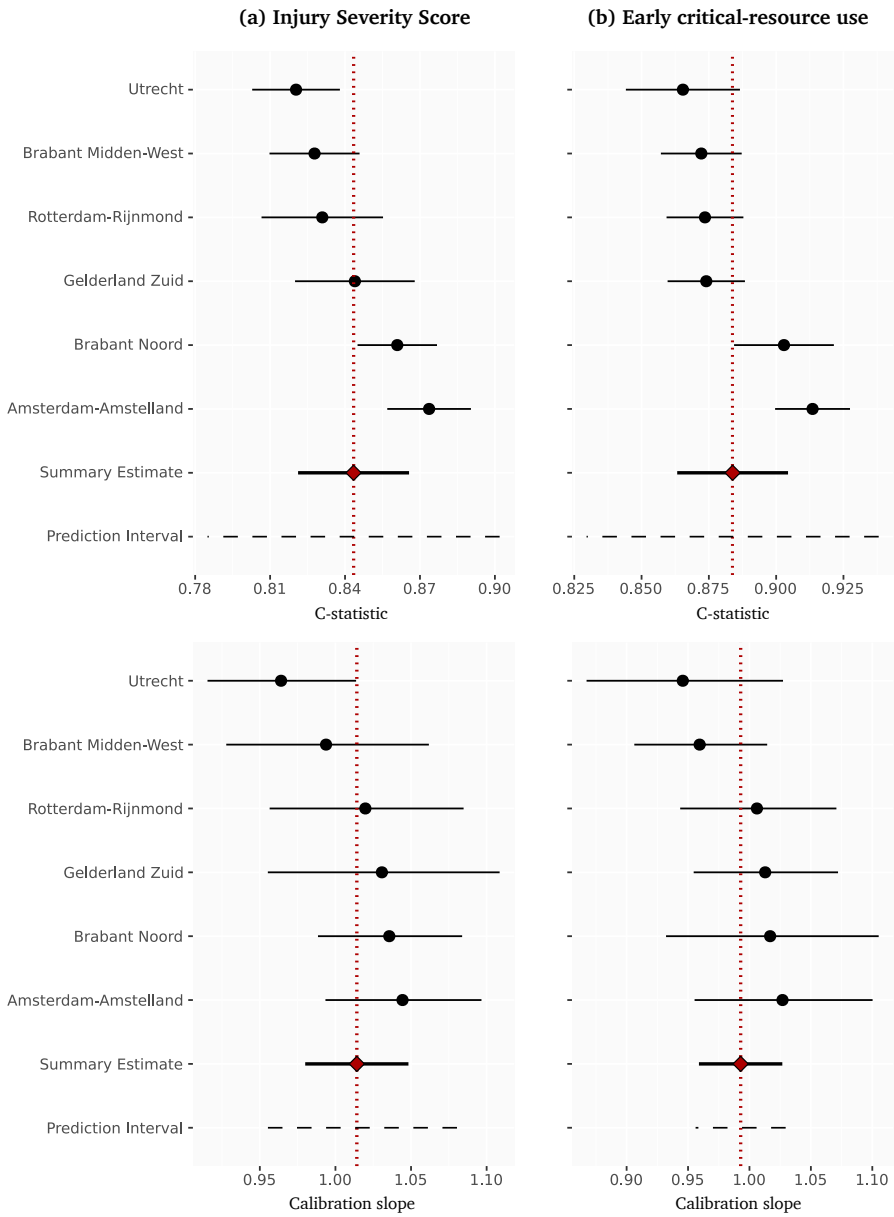


Figure 1 Univariate Meta-Analysis of Predictive Performance and Model Calibration

offset term.[18] Estimates of the recalibration intercept terms of the final models are presented per region to ease their implementation in practice.

Table 3 Results of Random-Effects Meta-Analyses for Model Performance

			EMS region					
Statistic	**Summary Estimate**	**95% (Approximate) Prediction Interval**	**Rotterdam-Rijnmond**	**Amsterdam-Amstelland**	**Brabant Midden-West**	**Brabant Noord**	**Utrecht**	**Gelderland Zuid**
Injury Severity Score ≥16								
C-statistic	0.84 (0.82 – 0.87)	0.79 – 0.90	0.86 (0.84 – 0.88)	0.87 (0.86 – 0.89)	0.82 (0.80 – 0.84)	0.84 (0.82 – 0.87)	0.83 (0.81 – 0.85)	0.83 (0.81 – 0.85)
Calibration intercept	-0.02 (-0.44 – 0.39)	-1.18 – 1.13	-0.27 (-0.42 – -0.12)	-0.58 (-0.74 – -0.43)	-0.02 (-0.20 – 0.15)	-0.08 (-0.31 – 0.14)	0.47 (0.20 – 0.75)	0.40 (0.17 – 0.63)
O:E ratio	1.00 (0.69 – 1.30)	0.13 – 1.86	0.77 (0.72 – 0.81)	0.59 (0.55 – 0.64)	1.07 (1.00 – 1.14)	0.95 (0.87 – 1.04)	1.32 (1.20 – 1.44)	1.30 (1.21 – 1.40)
Calibration slope	1.01 (0.98 – 1.05)	0.95 – 1.08	1.04 (0.99 – 1.08)	1.04 (0.99 – 1.10)	0.96 (0.92 – 1.01)	0.99 (0.93 – 1.06)	1.03 (0.96 – 1.11)	1.02 (0.96 – 1.08)
Recalibration intercept*	-0.02		-0.27	-0.54	0.08	-0.07	0.25	0.40
Early critical-resource use								
C-statistic	0.88 (0.86 – 0.90)	0.83 – 0.94	0.87 (0.86 – 0.89)	0.91 (0.90 – 0.93)	0.87 (0.86 – 0.89)	0.90 (0.88 – 0.92)	0.87 (0.84 – 0.89)	0.87 (0.86 – 0.89)
Calibration intercept	-0.06 (-0.62 – 0.51)	-1.63 – 1.51	-0.37 (-0.56 – -0.18)	-0.81 (-1.01 – -0.60)	0.22 (0.00 – 0.43)	-0.35 (-0.63 – -0.07)	0.44 (0.12 – 0.77)	0.56 (0.29 – 0.83)
O:E ratio	1.01 (0.74 – 1.28)	0.25 – 1.77	0.87 (0.83 – 0.92)	0.64 (0.60 – 0.67)	1.11 (1.05 – 1.18)	0.90 (0.83 – 0.97)	1.24 (1.15 – 1.33)	1.31 (1.24 – 1.39)
Calibration slope	0.99 (0.96 – 1.03)	0.96 – 1.03	0.96 (0.91 – 1.01)	1.01 (0.94 – 1.07)	1.01 (0.95 – 1.07)	0.95 (0.87 – 1.03)	1.02 (0.93 – 1.10)	1.03 (0.95 – 1.10)
Recalibration intercept*	-0.03		-0.15	-0.68	0.16	-0.16	0.22	0.42

Abbreviations: EMS, *Emergency Medical Service*; O:E, *Observed/Expected*. *Intercept of the simple logistic regression models used to recalibrate the initial predictions. Recalibration intercepts are from the final models and are therefore slightly different from the calibration intercepts.

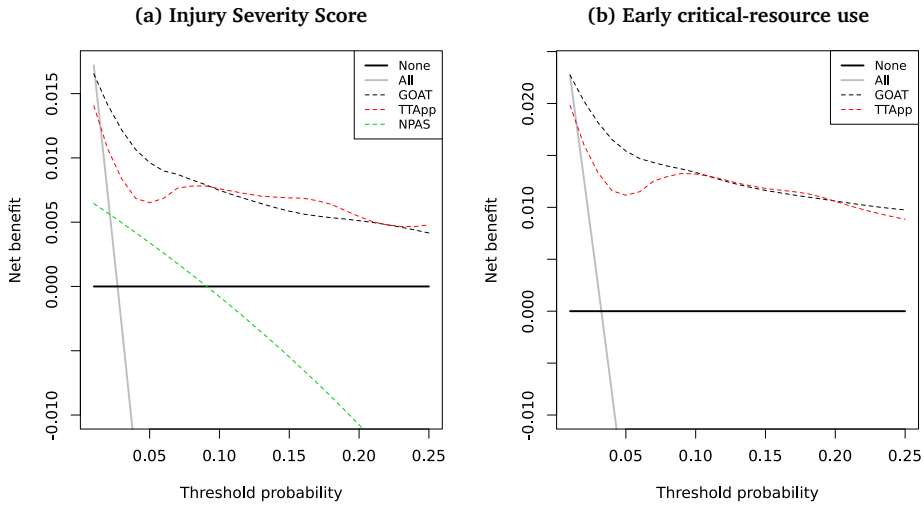


Figure 2 Decision Curve Analysis

Decision curve analyses of triage tools are presented in Figure 2. The relevant threshold interval was limited to 25%, as preventing undertriage is often preferred over low overtriage rates with a suggested factor of 5 to 7.[3] The novel models both have a superior net benefit for risk thresholds below 10%, and similar net benefit for higher thresholds, compared to the Trauma Triage App. The new models are superior to the National Protocol of Ambulance Services on the complete range of risk thresholds.
The c-statistic of the Trauma Triage App in the subsample (n=3629) was 0.77 (95%-CI, 0.71 – 0.84) based on the primary outcome. The c-statistic of the newly developed model was 0.07 (absolute increase of 14%) higher compared to the Trauma Triage App. Similar results were reported for the secondary outcome: a c-statistic of 0.81 (95%-CI, 0.75 – 0.86) and a difference of 0.07 (14%) in favor of the new model.

A table with diagnostic accuracy metrics for a wide range of probability thresholds is presented in the appendix.

DISCUSSION

In this prospective, multi-site study, we developed and externally validated two prediction models to select patients in need of specialized trauma care, based on IPD of 133,196 patients from six regions, 55 hospitals, and seven inclusive trauma regions. The developed models have been extensively validated, do not require complete information on predictor

variables, and require minimal human input. Random-effects meta-analyses were performed to investigate model performance and to identify between-region heterogeneity. We found that the newly developed models were highly accurate and model performance surpassed all triage tools used in contemporary practice in the Netherlands.

Comparison to Other Studies

Various tools have been proposed to aid pre-hospital trauma triage in the last decades.[19] The Trauma Score, Triage Revised Trauma Score, Pre-hospital Index, and CRAMS scale are point-based scoring systems developed in the 20th century.[7,20-22] These tools were all innovative at the time, showed improved accuracy compared to their predecessors, but have limited overall performance, local accuracy, or usefulness on-scene.[19,23] Nonetheless, these tools provide EMS professionals with categorical estimates of injury severity. In contrast, most present-day triage tools are decision schemes with dichotomous outcomes (e.g., transport to a higher-level trauma center or not).[3,6] Furthermore, none of these tools are both sensitive and specific, and model performance in different settings remains elusive. The Trauma Triage App was developed to improve predictive performance and was externally validated in a different (yet single) EMS region.[8] This tool is able to make predictions on patient-level but still requires substantial human input, non-missing predictor variables, and nationwide model performance remained unexplored.

We now provide a new model that outperforms contemporary used tools by a substantial margin and a new model to predict resource use. The decision curve analyses demonstrate superior net benefit of the new models for threshold probabilities below 10%, and similar benefit for higher risk thresholds. This implies that the predictions of the proposed models lead to better classifications than contemporary triage tools, and may thus improve medical decision making. Finally, our approach facilitates precision-medicine in the pre-hospital setting by providing EMS professionals with well-calibrated probabilities.[24]

Strengths and Limitations

This study has several strengths. The prospective, multi-site, Trauma Triage Continuum of Care Cohort was specifically designed for this study, included a consecutive series of patients suspected of injuries by EMS professionals during field triage, and was sufficiently powered to develop and validate complex prediction models.[9] To our knowledge, it is the largest, prospective, multi-center cohort designed to develop and validate prediction models aimed at optimizing pre-hospital trauma triage. A restricted set of simple, routinely collected, and automatically measured predictors was chosen to allow easy replication and uptake of the resulting models.

This is the first prediction model in pre-hospital trauma triage that rigorously assessed discrimination and model calibration across urban, suburban, and rural populations. Heterogeneity in model performance was analyzed thoroughly and intercepts to recalibrate predictions for new populations were reported. Heterogeneity was present in outcomes and in trauma care (e.g., length of stay), but there was very little difference in the case-mix of patient characteristics between regions. For this reason, the developed models are likely to generalize well across different trauma regions in the Netherlands, provided that aforementioned intercept updates are applied locally.

Possibly the biggest advantage of the models described in this investigation is their usability in time-critical situations. Implementation of these models does not require EMS professionals to deviate from routine care. In fact, predictions can be automatically generated based on electronic health records in real-time. The key difference between the prediction models developed in this study, other prediction models, and traditional triage tools is their ability to deal with missing values by design. This is of critical importance as missing data are inevitable in the pre-hospital setting. Model predictions can be generated in the absence of complete data and may be updated as new data become available.

A number of limitations should be addressed before interpreting the study results. First, the primary outcome is only moderately correlated with early critical-resource use.[25] An ISS of 16 or greater is often used as a surrogate marker to evaluate triage accuracy but was originally developed to predict mortality.[12] The secondary outcome, on the contrary, is correlated to resource-need by design, but is dependent on the initial transportation destination (i.e., resources might not have been used because they were unavailable). Second, we were unable to include a consecutive series of patients in 2018 (as was originally outlined in the study design) because of nationwide strikes in ambulance care, and therefore decided to shorten the inclusion period. This did not impact the appropriateness of the IPD, as each region still had a sufficient number of events.[26] Third, computational constraints limited the number of folds, iterations, and search space of the hyperparameter optimization procedure. Finally, it was not feasible to conduct random-effect meta-analyses to assess the clinical utility of the Trauma Triage App because of the labor-intensive methods needed to engineer predictors (e.g., suspected injury types) from unstructured text fields for over one hundred thousand patients. This is illustrative of the difficulties that may arise whenever models are based on predictors that are not routinely collected in daily practice.

Implications for Emergency Medical Services and Policy Makers

Undertriage is worldwide problem associated with potentially avoidable mortality.[1,4] Triage tools are an important component of the diagnostic or prognostic strategy to get the right patient to the right hospital. The initial transportation destination is mainly based on a tool's advice in conjunction with trauma center proximity, patient acuity, and the judgement of EMS professionals. Accurate models to predict need of specialized trauma care have large potential to lower mistriage rates. Our models may be used as decision-support tools at the scene of injury, could be incorporated into electronic health records, and could be deployed as Early Warning Systems. We argue that the model based on resource use will be more suited to use in inclusive trauma systems.

The prediction models (supplementary content) developed in this study can be implemented in digital decision support systems. In particular, these prediction models could be implemented into the software platform of the Trauma Triage App that is currently evaluated in daily practice in a nationwide stepped-wedge cluster randomized trial.[27] Ultimately, we aim to replace the National Protocol of Ambulance Services with the resource-based prediction model.

Further Research

Future research should focus on international validation and impact assessment of the proposed prediction models. Predictive accuracy might be improved by considering other variables, such as unstructured text fields and audio recordings. Deep neural networks (e.g., long-short term memory models) show promising results in various natural language processing tasks and might be well suited for pre-hospital electronic health records.[28] A different consideration might be to develop specialized tools to predict individual resource-need. A dynamic supply and demand model, based on personalized resource-need prediction, would finally unlock the true potential of inclusive trauma systems.

Acknowledgements

TPAD gratefully acknowledges financial support from the Netherlands Organization for Health Research and Development. We also like to thank all participating research sites, the Dutch Network of Acute Care, and Jill Whittaker for her advice on writing style.

Data Sharing

Our prediction models are freely available and can be found in the online supplemental content. Models are exported as JSON files with model parameters and binaries that can be loaded by the XGBoost packages in various programming environments, such as R, Julia, C/C++, and Python.

Role of the Funding Source
The funders of the study had no role in the study design, data collection, analyses, interpretation of the data, or the writing of the final report. All authors, internal and external, were able to access the full data on the primary research site and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Authors' Contributions:
RvdS, TPAD, LPHL, MP, and MvH contributed to the study design. RvdS, JFW, and RDL collected the data. RvdS, TPAD, and MvH analyzed the data. RvdS, TPAD, JFW, RDL, LPHL, MP, and MvH interpreted the data. RvdS, TPAD, and MvH drafted the manuscript. RvdS, TPAD, JFW, RDL, LPHL, MP, and MvH critically evaluated and revised the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no other meeting the criteria have been omitted. The publication is the work of the authors who will serve as guarantors for the contents of this paper.

REFERENCES

1. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

2. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

3. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Dec 20, 2019).

4. van Rein EAJ, van der Sluijs R, Houwert RM, et al. **Effectiveness of prehospital trauma triage systems in selecting severely injured patients: Is comparative analysis possible?** Am J Emerg Med 2018; 36(6): 1060-9.

5. Newgard CD, Fu R, Zive D, et al. **Prospective validation of the national field triage guidelines for identifying seriously injured persons.** J Am Coll Surg 2016; 222(2): 146-58 e2.

6. Voskens FJ, van Rein EAJ, van der Sluijs R, et al. **Accuracy of prehospital triage in selecting severely injured trauma patients.** JAMA Surg 2018; 153(4): 322-7.

7. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. **A revision of the Trauma Score.** J Trauma 1989; 29(5): 623-9.

8. van Rein EAJ, van der Sluijs R, Voskens FJ, et al. **Development and validation of a prediction model for prehospital triage of trauma patients.** JAMA Surg 2019.

9. van der Sluijs R, Debray TPA, Poeze M, Leenen LPH, van Heijl M. **Development and validation of a novel prediction model to identify patients in need of specialized trauma care during field triage: design and rationale of the GOAT study.** Diagn Progn Res 2019; 3: 12.

10. Collins GS, Reitsma JB, Altman DG, Moons KG. **Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement.** BMJ 2015; 350: g7594.

11. Riley RD, Ensor J, Snell KI, et al. **External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges.** BMJ 2016; 353: i3140.

12. 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(3): 187-96.

13. Gennarelli T, Wodzin E. **The Abbreviated Injury Scale 2005: update 2008**. Association for the Advancement of Automotive Medicine; Des Plaines, IL: 2008.

14. Chen T, Guestrin C. **XGBoost: A scalable tree boosting system.** arXiv:160302754 2016.

15. Audigier V, Resche-Rigon M. **micemd: multiple imputation by chained equations with multilevel data.** 2019. https://rdrr.io/cran/micemd/man/micemd-package.html (accessed Oct 9, 2019).

16. Debray TP, Moons KG, Ahmed I, Koffijberg H, Riley RD. **A framework for developing, implementing, and evaluating clinical prediction models in an individual**

participant data meta-analysis. Stat Med 2013; 32(18): 3158-80.

17. Rubin DB. **Multiple imputation for nonresponse in surveys**. New York: John Wiley & Sons, 1987.

18. Caruana R, Niculescu-Mizil A. **Predicting good probabilities with supervised learning.** Proceedings of the 22nd International Conference on Machine Learning 2005.

19. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld RA, Leenen LPH, van Heijl M. **Accuracy of prehospital triage protocols in selecting severely injured patients: a systematic review.** J Trauma Acute Care Surg 2017; 83(2): 328-39.

20. Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. **Trauma score.** Crit Care Med 1981; 9(9): 672-6.

21. Gormican SP. **CRAMS scale: field triage of trauma victims.** Ann Emerg Med 1982; 11(3): 132-5.

22. Koehler JJ, Baer LJ, Malafa SA, Meindertsma MS, Navitskas NR, Huizenga JE. **Prehospital Index: a scoring system for field triage of trauma victims.** Ann Emerg Med 1986; 15(2): 178-82.

23. Sewalt CA, Venema E, Wiegers EJA, et al. **Trauma models to identify major trauma and mortality in the prehospital setting.** Br J Surg 2019.

24. Jameson JL, Longo DL. **Precision medicine--personalized, problematic, and promising.** N Engl J Med 2015; 372(23): 2229-34.

25. Newgard CD, Hedges JR, Diggs B, Mullins RJ. **Establishing the need for trauma center care: anatomic injury or resource use?** Prehosp Emerg Care 2008; 12(4): 451-8.

26. Collins GS, Ogundimu EO, Altman DG. **Sample size considerations for the external validation of a multivariable prognostic model: a resampling study.** Stat Med 2016; 35(2): 214-26.

27. van der Sluijs R, Fiddelers AAA, Waalwijk JF, et al. **The impact of the Trauma Triage App on prehospital trauma triage: design of the stepped-wedge, cluster randomized TESLA trial.** Available from: https://www.trialregister.nl/trial/7038 (accessed Dec 9, 2019).

28. Devlin J, Chang MW, Lee K, Toutanova K. **BERT: Pre-training of deep bidirectional transformers for language understanding.** arXiv:181004805 2018.

APPENDIX 1. HOW TO USE THE MODELS

Two models are provided in the supplementary content: (i) a model to predict injury severity (*outcome1.model*) and (ii) a model to predict in-hospital resource use (*outcome2.model*). These models can be loaded by the xgboost packages in several programming languages (e.g., Python, R, Julia). Predictions on patient-level can be generated using three simple steps: (i) set the predictor values; (ii) load the preferred model; and (iii) finally generate predictions.

```
# Packages
library(xgboost)

# 1) Create a dummy matrix with predictor values for 10 patients
newdata =
  tibble(
    Leeftijd = rnorm(10, 61.5, 15), # Age in years
    Man = round(runif(10)), # Male (1 = yes, 0 = no)
    Weekdag = ceil(runif(10, 0, 7)), # Weekday (Monday = 1, Sunday = 7)
    Uur = ceil(runif(10, 0, 24)), # Hour of the day
    RT = rnorm(10, 8.9, 3), # Response time
    Assistentie_MMT = rbinom(10, 1, 0.025), # Medical doctor assistance on-scene
    SBP_first = round(rnorm(10, 140, 15)), # Systolic blood pressure
    DBP_first = round(rnorm(10, 90, 10)), # Diastolic blood pressure
    E_first = ceil(runif(10, 0, 4)), # Eyes component of the GCS
    M_first = ceil(runif(10, 0, 6)), # Motor component of the GCS
    V_first = ceil(runif(10, 0, 5)), # Verbal component of the GCS
    HR_first = round(rnorm(10, 81, 25)), # Heart rate
    RR_first = round(rnorm(10, 16, 4)), # Respiratory rate
    SPO2_first = round(rnorm(10, 97, 1.5)), # Oxygen saturation
    B_Intubatie = rbinom(10, 1, 0.023), # Out-of-hospital intubation
    MKA_Urgentie_A1 = rbinom(10, 1, 0.51) # Highest ambulance priority
  ) %>%
  mutate(
    GCS_first = E_first + M_first + V_first # Glasgow Coma Scale
  ) %>%
  as.matrix

# 2) Load one of both models
fit <- xgb.load("outcome1.model")

# 3) Predict
predict(fit, newdata)
```

APPENDIX 2. SUPPLEMENTARY TABLES

Table S1 Definition of Major Surgical Intervention

Damage control thoracotomy
Damage control laparotomy
Extra peritoneal pelvic packing
Revascularization of extremities
Craniotomy
Coniotomy/cricothyrotomy
Damage control orthopedics

One or more of these interventions result in a positive secondary outcome (i.e., resource use was considered to be positive).

Table S2 List of Threshold Probabilities and Associated Diagnostic Accuracy

Threshold*	Sensitivity	Specificity	Undertriage*	Overtriage*
Injury Severity Score ≥16				
0.005	0.92	0.46	0.08	0.54
0.0075	0.90	0.53	0.10	0.47
0.01	0.84	0.65	0.16	0.35
0.02	0.71	0.85	0.29	0.15
0.03	0.65	0.91	0.35	0.07
0.04	0.61	0.93	0.39	0.07
0.05	0.56	0.95	0.44	0.05
0.10	0.48	0.97	0.52	0.03
0.15	0.39	0.99	0.61	0.01
Early critical-resource use				
0.005	0.95	0.45	0.05	0.55
0.01	0.93	0.54	0.07	0.46
0.02	0.82	0.80	0.18	0.20
0.03	0.72	0.92	0.28	0.08
0.04	0.69	0.93	0.31	0.07
0.05	0.66	0.95	0.34	0.05
0.10	0.60	0.98	0.40	0.02
0.15	0.51	0.99	0.49	0.01

*Cases are considered positive (i.e., in need of specialized trauma care) if the predicted probability is equal to or greater than the threshold. *Undertriage and overtriage metrics in a hypothetical situation with the prediction model as the only diagnostic criterion and complete adherence to its advice.

Table S3 Software Packages and Version Information

Package	Version
Python	3.7
xgboost	0.82
scikit-learn	0.21.3
hyperopt	0.2.1
numpy	1.17.2
matplotlib	3.1.1
tqdm	4.36.1
pandas	0.25.2
R	3.6.1
tidyverse	1.2.1
CalibrationCurves	0.1.2
pROC	1.15.3
boot	1.3-23
metamisc	0.2.2
mice	3.6.0
micemd	1.6.0
lattice	0.20-38
gridExtra	2.3
Amelia	1.7.6

APPENDIX 3. SUPPLEMENTARY FIGURES

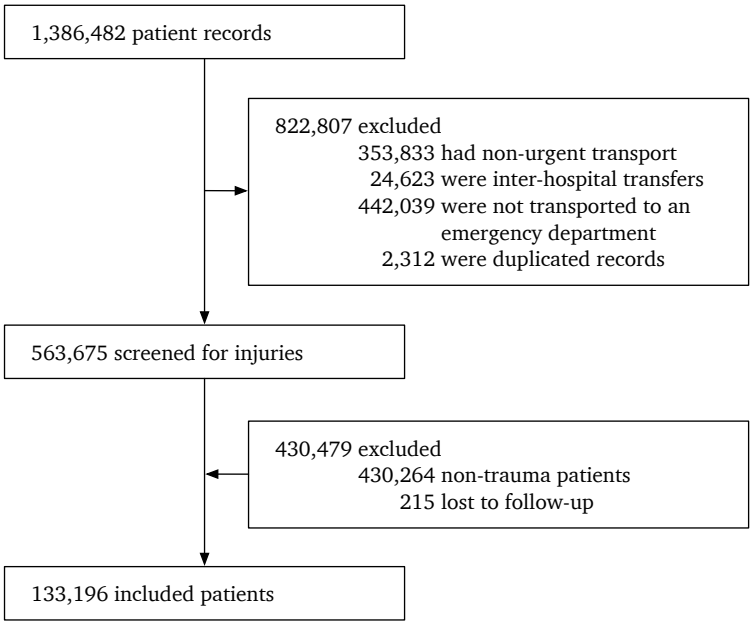


Figure 1 Study Profile

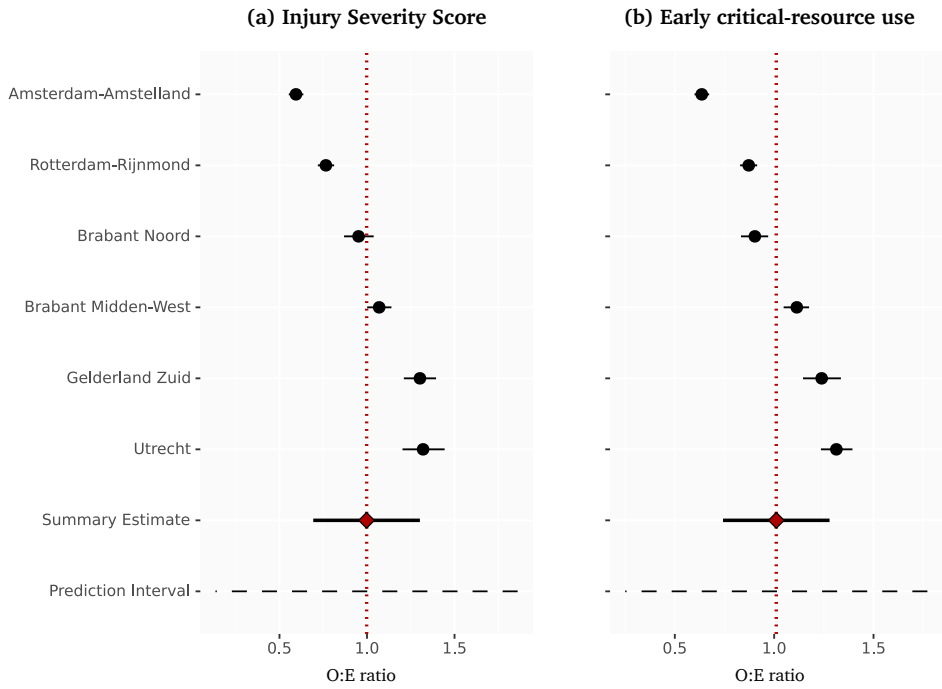
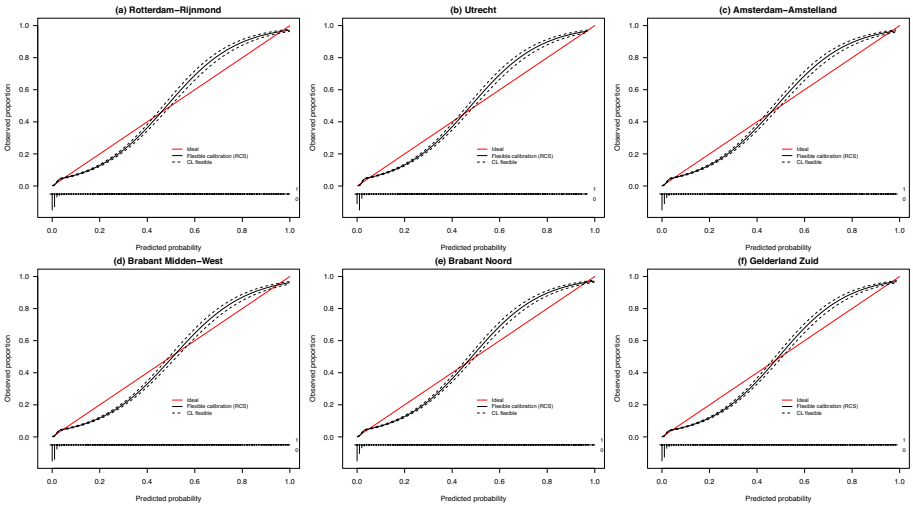


Figure 2 Univariate Meta-Analyses of Observed versus Expected Ratios

(a) Injury Severity Score ≥16



(b) Early Critical-Resource Use

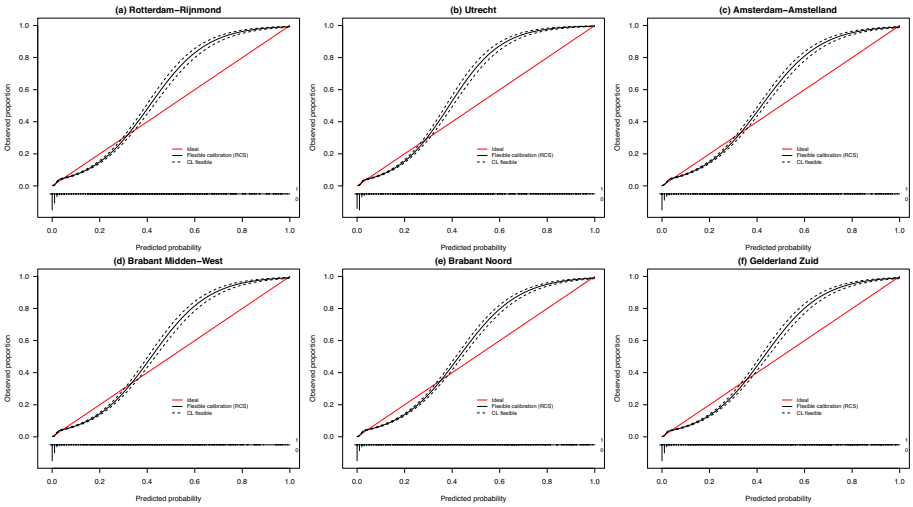


Figure 3 Calibration Plots of Observed Proportions Versus Predicted Probabilities

CHAPTER IX

The impact of the Trauma Triage App on pre-hospital trauma triage: design of the stepped-wedge, cluster randomized TESLA trial

Diagnostic & Prognostic Research

Rogier van der Sluijs; Audrey A.A. Fiddelers; Job F. Waalwijk; Johannes B. Reitsma; Miranda J. Dirx; Dennis den Hartog; Silvia M.A.A. Evers; J. Carel Goslings; W. Margreet Hoogeveen; Koen W. Lansink; Luke P.H. Leenen; Mark van Heijl; Martijn Poeze

ABSTRACT

Background

Field triage of trauma patients is crucial to get the right patient to the right hospital within a particular time frame. Minimization of undertriage, overtriage and inter-hospital transfer rates could substantially reduce mortality rates, life-long disabilities and costs. Identification of patients in need of specialized trauma is predominantly based on judgment of Emergency Medical Services professionals and a pre-hospital triage protocol. The Trauma Triage App is a smartphone application that includes a prediction model to aid Emergency Medical Services professionals in the identification of patients in need of specialized trauma care. The aim of this trial is to assess the impact of this new digital approach to field triage on the primary endpoint undertriage.

Methods

The Trauma triage using Supervised Learning Algorithms (TESLA) trial is a stepped-wedge cluster randomized controlled trial with eight clusters defined as Emergency Medical Services regions. These clusters are an integral part of five inclusive trauma regions. Injured patients, evaluated on-scene by Emergency Medical Services professionals, suspected of moderate to severe injuries, will be assessed for eligibility. This unidirectional crossover trial will start with a baseline period in which the default pre-hospital triage protocol is used, after which all clusters gradually implement the Trauma Triage App as an add-on to the existing triage protocol. The primary endpoint is undertriage on patient and cluster level and is defined as the transportation of a severely injured patient (Injury Severity Score ≥16) to a lower-level trauma center. Secondary endpoints include overtriage, hospital resource use and a cost-utility analysis.

Discussion

The TESLA trial will assess the impact of the Trauma Triage App in clinical practice. This novel approach to field triage will give new and previously undiscovered insights into several isolated components of the diagnostic strategy to get the right trauma patient to the right hospital. The stepped-wedge design allows for within and between cluster comparisons.

BACKGROUND

Pre-hospital trauma triage is crucial to match an injured patient to the optimal definitive care facility.[1] Erroneously transporting a patient requiring specialized trauma care to a lower-level trauma center, could lead to a delay in definitive care and is associated with higher mortality and morbidity rates.[2] Conversely, transporting a patient not in need of specialized trauma care to a higher-level trauma center, results in extra costs and overutilization of resources.[3] These key metrics to evaluate the quality of field triage in trauma systems are termed undertriage and overtriage, respectively.[1] The Dutch National Health Care Institute guidelines state that a maximum of 10% undertriage is acceptable in the Netherlands.[4] The mean undertriage across all inclusive trauma regions in the Netherlands was 31.4% in 2016.[5]

The first step of the multivariable strategy to determine the optimal receiving facility is to identify patients in need of specialized trauma care. This is performed by Emergency Medical Services (EMS) professionals on-scene and is influenced by both the pre-hospital triage protocol and the judgment of the EMS professional. The 8th version of the National Protocol of Ambulance Services (NPAS; in Dutch, *Landelijk Protocol Ambulancezorg*), is currently used by all EMSs in the Netherlands as the primary pre-hospital triage protocol. A recent study reported that strict adherence to the criteria of the NPAS would have led to an undertriage rate of 63.8%, with an overtriage rate of 7.3% in one inclusive trauma region.[6] Moreover, a systematic review did not identify any pre-hospital triage protocol that by itself complied with the target of 10% undertriage.[7]

The Trauma Triage App (TTApp), a smartphone and tablet application that incorporates a prediction model, was recently developed to identify patients in need of specialized trauma care. The main function of the TTApp is to predict an individual patient's probability of being severely injured. An advice regarding whether a patient requires specialized trauma care is then generated based on a pre-defined threshold probability. This novel approach to trauma triage was externally validated retrospectively in 6859 patients from a different EMS and was able to retain a c-statistic of 0.83 with proper calibration.[8] An undertriage rate of 11.2% with a combined overtriage rate of less than 50% can hereby be achieved, depending on the threshold probability.

Notwithstanding these promising validation results, the impact of the use of the TTApp in daily practice remains to be established. The TESLA (Trauma Triage using Supervised Learning Algorithms) stepped-wedge cluster randomized trial was designed to evaluate whether the availability of the TTApp during field triage indeed leads to a decrease in undertriage, while preserving acceptable overtriage rates.

METHODS/DESIGN

Study Design

The TESLA trial is a prospective, stepped-wedge cluster randomized trial (SW-CRT). In a SW-CRT, clusters are randomized into allocation sequences. These sequences all start with one or more periods under the control condition, followed by the remaining periods in which the intervention is implemented. In this trial, the participating EMS regions (the clusters) will be randomized upfront to determine the period after which two paired clusters will switch to the intervention condition. The allocation sequence will be defined using computer-generated random numbers on the primary research site.

Our aim is to include 1920 consecutive severely injured patients in five steps, each with a duration of four months (details about the sample size below). All clusters will start with one or more steps of usual care (the NPAS). At the end of each step, two clusters will switch from the NPAS to the intervention condition (the TTApp used as an add-on to the NPAS). Key study design features are shown in Figure 1.

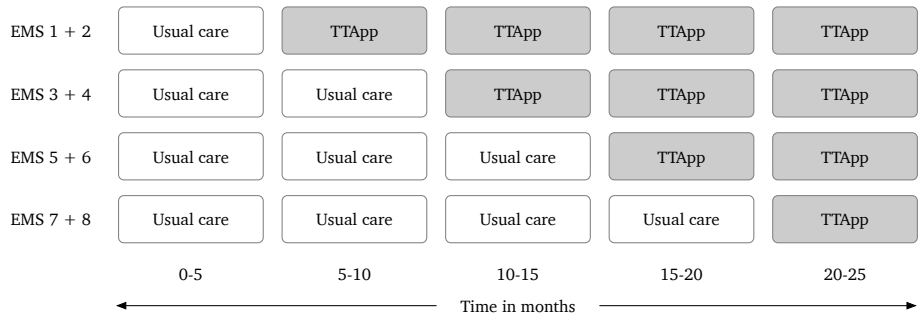


Figure 1 Timing of Implementation of the Intervention for Each Pair of Emergency Medical Services. Abbreviations: EMS, *Emergency Medical Service*; TTApp, *Trauma Triage App*.

Participating Regions

Eight out of 25 EMSs (the clusters) in the Netherlands, with approximately equal patient volumes, were selected to participate in this trial (Figure 2). These EMSs are an integral part of five distinct inclusive trauma regions and cover urban, suburban and rural areas. All 37 hospitals with a trauma-receiving emergency department within these five distinct inclusive trauma regions participate in the collection of relevant patient outcomes. All hospitals with an emergency department within these regions are designated a level of care being either one, two or three. Level-I is considered a higher-level trauma center, whereas level-II and III trauma centers are acknowledged as lower-level trauma centers.

Figure 2 Service Regions of the Participating Emergency Medical Services

Study Population

All patients, 18 years of age or older, evaluated on-scene by an EMS professional, suspected of moderate to severe injuries, defined as an Injury Severity Score (ISS) of 9 or greater, will be assessed for eligibility. Patients transported to a hospital outside of the participating trauma regions will be excluded. Patients that are death on arrival at the initial receiving emergency department will also be excluded.

Pre-hospital Trauma Triage Tools

Usual Care: National Protocol of Ambulance Services

No adjustments to daily practice of EMS professionals will be introduced prior to the switch to the intervention period. Assessment of injury severity is often a two-step process consisting of the evaluation of the pre-hospital trauma triage protocol and the final judgment of the EMS professional. The 8th version of the NPAS is currently used by all participating EMSs. This protocol is a flowchart that consists of multiple criteria to

identify patients in need of specialized trauma care. If a patient fulfills one of the criteria in Table 1, the EMS professional is advised to transport the patient to a higher-level trauma center. In daily practice, it is not obligatory to adhere to the protocol or to report its advice.

Table 1 Higher-Level Trauma Center Criteria of the National Protocol of Ambulance Services

ABC-unstable during evaluation on-scene
Revised Trauma Score <11
Deteriorating Glasgow Coma Scale score
Glasgow Coma Scale score <9
Flail chest
Amputation proximal to wrist or ankle
Two or more fractures of femur and/or humerus
Penetrating injury of head, thorax or abdomen
Unstable pelvic fracture
Body temperature <32 degrees Celsius
Neurologic deficit of one or more extremities
Anisocoria

Intervention: Trauma Triage App

The TTApp is a smartphone and tablet application for both Android and iOS operating systems (Figure 3). The application is a practical and quick-to-use questionnaire consisting of six questions that collect the required predictor values of the prediction model incorporated in the TTApp (Table 2). Four additional questions are added: the judgment of the EMS professional prior to the questionnaire, the judgment of the EMS professional after the advice returned by the prediction model was given, the transportation destination while displaying a map with distances to nearby hospitals, and when applicable, a screen to specify reasons to bypass the preferred hospital. The incorporated prediction model calculates the probability that a patient is severely injured. An advice whether to transport a patient to a higher-level trauma center or not is generated based on a pre-defined threshold probability. This threshold determines the sensitivity and specificity of the prediction model. Filling out the questionnaire takes approximately 30–45 seconds and should be performed on-scene by EMS professionals.

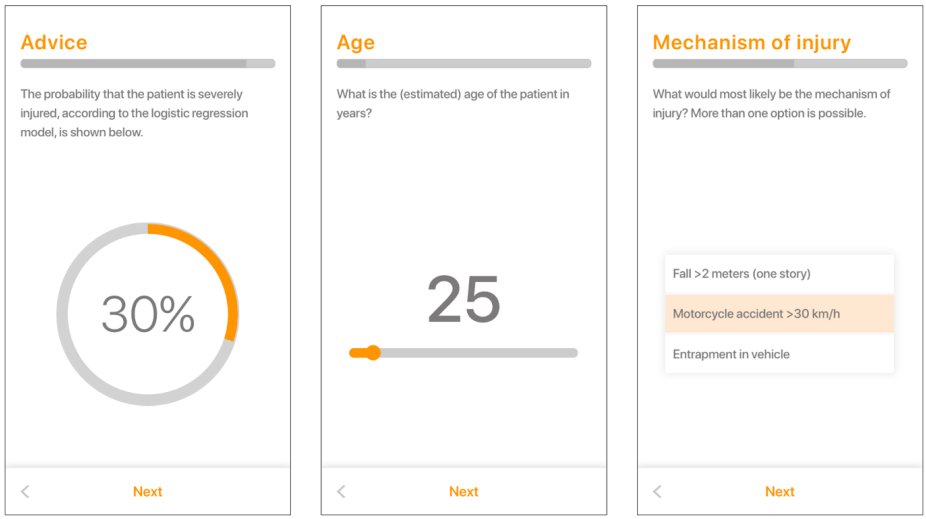


Figure 3 A Sample of Screens from the Trauma Triage App. Left: the generated score indicating the probability that a patient might be severely injured based on all predictors. Middle: an input field requesting the age of the patient in years. Right: an input field requesting the mechanism of injury.

Decision-making about whether to transport a patient to a higher or lower-level trauma center will be carried out similarly to usual care, with the exception of the availability of the TTApp prediction and recommendations linked to that prediction. Also, alike the NPAS, the TTApp is a decision-support system that can be overruled by EMS professional judgment. Implementation of the TTApp will likely lead to the transportation of more (severely) injured patients to higher-level trauma centers, thus reducing undertriage. This might lead to a slight increase in overtriage.

The TTApp will be introduced to all EMSs in a systematic manner at the end of the baseline period. A presentation will be provided to all EMS professionals that teaches the rationale behind pre-hospital triage and the study protocol. An electronic-learning will be made available that will demonstrate the use of the TTApp. The application will subsequently be made available on the proprietary devices of the participating EMS.

Data Collection

Pre-hospital data of trauma patients is routinely collected by EMS professionals through digital run-reports. Extra variables, mainly answers to the questionnaire and usage information, will be collected by the TTApp in the intervention period. Clinical data, including all relevant patient outcomes, of all patients admitted to one of the participating hospitals will be collected through the trauma registries of each inclusive trauma region.

Table 2 Variables of the Prediction Model Incorporated in the Trauma Triage App

Age
Systolic blood pressure, mm Hg
Glasgow Coma Scale score
Penetrating injury of head, thorax or abdomen
Fall >2 m or motorcycle accident >30 km/h or entrapment in vehicle
Suspected moderate or severe head injury
Suspected moderate or severe thoracic injury
Injuries in at least two anatomical regions (head/neck, face, thorax, abdomen, extremities and/or external injuries

Injury Severity Scores will be routinely calculated within 30 days after the date of injury by trained trauma data managers for all admitted patients and those who die in the emergency department. All other included patients are assumed to have minor or moderate injuries (ISS <16). This assumption was validated for all patients discharged from the emergency department in a previous study.[6] Hospital data and pre-hospital data will be anonymized first, and then linked using a combined deterministic and probabilistic linkage scheme. This anonymized linkage approach was validated to be both highly sensitive and specific in prior research.[6] The final dataset will be accessible by JFW, MvH, and MP.

Patient Safety

The TTApp is a diagnostic intervention aimed at EMS professionals. Regular care by EMS professionals should not be impacted by the TTApp. The prediction model is a decision support tool, that – alike the NPAS – can be overruled by EMS professionals. This model is more sensitive and less specific compared to the NPAS, implicating that severely injured patients have a higher chance to get an advice for transportation to a higher-level trauma center. It will not be mandatory for EMS professionals to use the TTApp during the intervention period.

Primary Outcome

Primary endpoint of the study is undertriage, defined as the transportation of a severely injured patient (ISS ≥16) transported from the scene of injury to a lower-level trauma center. This implies that on patient-level, a severely injured patient can be either correctly triaged and transported to a higher-level trauma center or incorrectly triaged and transported to a lower-level trauma center.

Secondary Outcomes

- Overtriage, defined as the transportation of non-severely injured patients (ISS <16) from the scene of injury to a higher-level trauma center, will be evaluated on patient-level.
- A non-compliance analysis will be conducted to evaluate the efficacy of the TTApp under ideal circumstances with complete adherence by all EMS professionals in the specified study population.
- Use of health care resources. A comparative analysis will be performed to evaluate the differences in hospital length of stay, number of admissions to the Intensive Care Unit and length of stay at the Intensive Care Unit, between control and intervention conditions.
- The diagnostic accuracy of the prediction model incorporated in the TTApp will be evaluated for all eligible patients.
- A cost-effectiveness analysis will be performed alongside this SW-CRT that is described in a separate protocol.

Statistical Analysis

Primary Analysis

The primary endpoint – undertriage – will be analyzed at a patient-level using a generalized linear mixed model (GLMM). A random intercept will be introduced in the model to account for cluster differences. Time will be modelled as a categorical variable denoting the cluster step. The GLMM will be used in conjunction with the binomial distribution and the identity link, resulting in a risk difference between the control and intervention condition. Bootstrapped 95%-CIs will be estimated from this model. This intention-to-treat analysis will be adjusted for age, which is expected to be non-linear and thus will be modelled using restricted cubic regression splines. Missing values will be multiply imputed using a multi-level multiple imputation strategy that accounts for cluster differences.

Secondary Analyses

- Overtriage is analyzed using the same strategy as the primary analysis: a GLMM using a binomial distribution and the identity link resulting a risk difference. Bootstrapped 95%-CIs will be calculated, and analyses will be adjusted for age. All statistics are calculated for patients with an ISS <16.
- The primary analysis is aimed to assess the effectiveness of the TTApp. A non-compliance analysis, using instrumental variable estimation, will be conducted to evaluate the efficacy of the TTApp in the hypothetical situation with complete adherence.[9,10]
- Healthcare resources measured on a continuous scale (e.g., length of stay) are analyzed

using a GLMM with a Gaussian distribution and the identity link. Numbers of admissions are converted to proportions and analyzed similar to the primary analysis.

- The probability generated by the logistic regression model incorporated in the TTApp will be calculated for all eligible patients based on the digital run-reports and the data generated by the TTApp. Diagnostic accuracy measures, such as sensitivity, specificity, predictive values and likelihood ratios, will be calculated and model discrimination and calibration will be assessed.

Sample Size

The primary goal of the TESLA study is to evaluate whether implementation of the TTApp in daily practice reduces undertriage. The sample size calculation is based on this endpoint. The intra-cluster correlation was calculated using the Fleiss-Cuzick method and was 0.098.[11] The proportion of expected undertriage under usual care was 0.35, whereas a decrease of 0.1 was expected during the intervention period. With eight clusters, a power of 80%, a significance level of 0.05 and two clusters switching to the intervention after every step, at least 48 severely injured patients will have to be included per cluster per step.[12,13] Approximately 1300-1400 severely injured patients will be transported by the participating EMSs on yearly basis, therefore we expect a duration of (less than) four months per step. The total study time consists of a baseline period plus four additional steps, totaling 20 months in which 1920 severely injured patients should be included.

Approximately 1536 severely injured patients were included until Dec 1, 2019, of which 576 patients were recruited during the intervention period.

DISCUSSION

Getting the right patient to the right hospital within a certain time frame is becoming increasingly important with the maturation of trauma systems and centralization of resources. Costs and mortality rates can be reduced by minimizing undertriage, overtriage and inter-hospital transfer rates. The optimal hospital for an individual injured patient has to be determined on-scene by a diagnostic strategy that consists predominantly of (1) identification of injured patients in need of specialized trauma care and (2) logistical considerations, such as trauma center proximity and trauma center capacity. Pre-hospital triage tools, such as the Field Triage Decision Scheme, the NPAS, and the TTApp attempt to assist EMS professionals in the first step of this strategy. These diagnostic tools should be thoroughly tested and preferably evaluated by both external validation and impact assessment before widespread implementation in clinical practice. External validation is

crucial to evaluate the true performance of a prediction model in new (external) patients. Although the previously described prediction model is externally validated already, the application of the study results in practice might uncover implementation problems, disadvantages of a digital approach, possible improvements and might thus yield different results than expected. The aim of the TESLA trial is therefore to assess the impact of the TTApp in practice. This will likely give insights into reasons for nonadherence, reasons to overrule the prediction model and its advice, and the isolated impact of many of the components in the diagnostic strategy that lead to the determination of the most optimal hospital.

Stepped-wedge cluster randomized designs are particularly used to evaluate the impact of the implementation of prediction models in clinical practice.[14] This unidirectional crossover design combines elements of before-after studies with cluster randomization and is an efficient design that enables to derive a valid answer for the research question.

This study is limited by the fact that current population values were used to determine the sample size. These values might not reflect actual event rates during the trial. Patients require an ISS ≥9 to be eligible for inclusion in this trial. This is based on the assumption that EMS professionals are able to differentiate between mildly injured patients and those who are severely injured, which is likely, but might not entirely resemble actual usage. This constraint was posed to limit overtriage of a clearly non-severely injured group of patients. Another potential limitation of the trial is the innovativeness of the TTApp and subsequently its dissimilarity to routinely used static decision schemes. This could potentially lead to substantial nonadherence. Due to these reasons and because of the fixed length of the steps in a SW-CRT and its inextensible nature, the sample size was calculated with a conservative estimate of the expected decrease in undertriage.

CONCLUSIONS

The TESLA-trial is a SW-CRT that aims to evaluate the impact of the TTApp on the primary endpoint undertriage, as well as overtriage and hospital resource use. The smartphone application can potentially acquire new and previously undiscovered insights into several components of the strategy that leads to the determination of the optimal hospital for a specific injured patient.

Funding
This study is financially supported by The Netherlands Organization for Health Research and Development, Effectiveness Program (ZonMw [https://www.zonmw.nl/en/], dossier

number 80-84300-98-18555). The TESLA-trial is an investigator-initiated trial. The funding body has no role in the design of the study, collection, analysis, and interpretation of data and neither in writing the manuscript.

Authors' Contributions

RvdS contributed to the study design, wrote the statistical analysis plan, and drafted and revised the paper. AAAF, JFW, JBR, LPHL, MvH and MP contributed to the study design, analysis plan and revised the paper. All authors were involved in the study design and revised the draft paper.

REFERENCES

1. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Aug 20, 2019).

2. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

3. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

4. Zorginstituut Nederland. **Spoed moet goed: indicatoren en normen voor zes spoedzorgindicaties;** 2015.

5. Landelijk Netwerk Acute Zorg. **Landelijke Traumaregistratie 2012-2016: Rapportage Nederland.** Utrecht, 2017.

6. Voskens FJ, van Rein EAJ, van der Sluijs R, et al. **Accuracy of prehospital triage in selecting severely injured trauma patients.** JAMA Surg 2018; 153(4): 322-7.

7. van Rein EAJ, Houwert RM, Gunning AC, Lichtveld RA, Leenen LPH, van Heijl M. **Accuracy of prehospital triage protocols in selecting severely injured patients: A systematic review.** J Trauma Acute Care Surg 2017; 83(2): 328-39.

8. van Rein EAJ, van der Sluijs R, Voskens FJ, et al. **Development and validation of a prediction model for prehospital triage of trauma patients.** JAMA Surg 2019.

9. Hernan MA, Hernandez-Diaz S. **Beyond the intention-to-treat in comparative effectiveness research.** Clin Trials 2012; 9(1): 48-55.

10. Toh S, Hernan MA. **Causal inference from longitudinal studies with baseline randomization.** Int J Biostat 2008; 4(1): Article 22.

11. Wu S, Crespi CM, Wong WK. **Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials.** Contemp Clin Trials 2012; 33(5): 869-80.

12. Hooper R, Bourke L. **Cluster randomised trials with repeated cross sections: alternatives to parallel group designs.** BMJ 2015; 350: h2925.

13. Hooper R, Teerenstra S, de Hoop E, Eldridge S. **Sample size calculation for stepped wedge and other longitudinal cluster randomised trials.** Stat Med 2016; 35(26): 4718-28.

14. Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. **Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation.** J Clin Epidemiol 2011; 64(9): 936-48.

PART IV

Dynamic trauma systems

CHAPTER X

A personalized strategy to evaluate and optimize pre-hospital trauma systems

Rogier van der Sluijs; Luke P.H. Leenen; Mark van Heijl

ABSTRACT

Regionalized systems of trauma care are designed to centralize resources, expertise, and patients, in order to enable optimal and cost-efficient trauma care. Mistriage in inclusive trauma system must be averted since it is associated with increased mortality rates, exhaustive use of scarce resources, and excessive costs. The American College of Surgeons Committee on Trauma periodically publishes guidelines on evaluation of pre-hospital trauma systems and recommends undertriage rates of less than 5%, and to use the Field Triage Decision Scheme to determine the optimal transportation destination at the scene of injury, based on a crude assessment of injury severity (severely injured versus non-severely injured). The Injury Severity Score with a cut-off at 16 is the most widely used and recommended reference standard to evaluate field triage but is only moderately correlated with early critical-resource use. Composite resource-based reference standards are frequently proposed as a more valid alternative to anatomical reference standards but face new methodological challenges. This article reflects on flaws in current reference standards, triage tools, and the evaluation of pre-hospital trauma systems. We provide recommendations to make the field triage strategy transparent in order to identify to malfunctioning components (e.g., triage tools, EMS judgement). In addition, we illustrate the imperfections of surrogate reference standards and current resource-based proposals. Finally, we outline the design of a precision medicine approach to get the right patient to the right hospital.

INTRODUCTION

The premise of regionalized trauma systems is that centralization of patients and resources enables the most efficient trauma care. The importance of correct field triage and its impact on patient outcomes is proportionally related to the amount of centralization within such systems.[1,2] In this context, the objective of triage is to match injured patients to facilities with appropriate resources. This process may be perceived as a three-step diagnostic strategy in which Emergency Medical Services professionals first determine the resource-need of an injured patient, then evaluate logistical constraints (e.g., trauma center capacity, patient acuity and trauma center proximity), and ultimately define the optimal transportation destination in light of steps one and two.

Since field triage is the main entry to specialized care within trauma systems, mistriage could be disastrous. Undertriage – transporting patients to a facility without sufficient resources for optimal treatment – is a medical problem associated with increased mortality rates.[3] Conversely, overtriage – transporting patients to a facility with a surplus of resources for optimal treatment – results in excessive costs and overutilization of finite resources.[4] Extensive evaluation and subsequent refinement of the triage process is critical to avoid mistriage. The current evaluation methodology, however, leads to an inaccurate representation of clinical practice. This article reflects on this methodology and proposes a new, personalized, evaluation and optimization strategy.

EVALUATING THE PRE-HOSPITAL TRAUMA SYSTEM

Current guidelines by the American College of Surgeons Committee on Trauma (ACSCOT) state that trauma systems must aim to lower undertriage to less than 5%.[5] The ACSCOT suggests to dichotomize patients based on injury severity (severely-injured versus non-severely injured) and hospitals based on resources (higher-level versus lower-level trauma centers) in order to evaluate mistriage rates (i.e, the accuracy of the pre-hospital trauma system). Triage tools, such as the Field Triage Decision Scheme of the ACSCOT, also dichotomize patients based on simple criteria to determine the optimal transportation destination at the scene of injury.[5] Overall undertriage and overtriage rates can be derived from a contingency table displaying the frequency distributions of the aforementioned groups (Table).

There are several issues with this strategy. First, these metrics provide little information and are non-actionable. It is impossible to tell which components (e.g., triage tools, EMS judgement, logistical constraints, etc.) of the triage strategy function properly, and which

components could be improved based on overall undertriage and overtriage rates. Second, undertriage and overtriage are largely dependent on the baseline probability of being transported to a higher-level trauma center (e.g., undertriage is typically low in a system with a high density of higher-level trauma centers) and should thus be interpreted jointly. Third, specifying a target for undertriage, with varying amounts of acceptable overtriage, hinders comparison between trauma systems.

Dichotomization of patients and hospitals is practical, as it eases the pre-hospital decision process and its evaluation. The downside of this dichotomization is that it yields an unrealistic representation of the heterogeneous spectrum of trauma patients. This same principle applies to dividing the various levels of trauma centers into lower-level and higher-level centers. These abstractions will lead to suboptimal allocation of patients in inclusive trauma systems. In a more complex – real life – situation with dynamic availability of resources, it could even lead to malfunctioning trauma systems. In the era

Table 1 Example of Mistriage Rates Based on a Contingency Table

	Need for Specialized Trauma Care	
	Yes	**No**
Transported to higher-level TC		
Yes	850 (TP)	1500 (FP)
Resource 1	800	1000
Resource 2	50	500
No	150 (FN)	3000 (TN)
Resource 1	100	1500
Resource 2	50	1500
Mistriage rates	**Estimate**	**95%-CI***
Undertriage: FN / (TP + FN)	15%	13% - 17%
Resource 1	11%	9% - 13%
Resource 2	50%	40% - 60%
Overtriage: FP / (TN + FP)	33%	32% - 35%
Resource 1	40%	38% - 42%
Resource 2	25%	23% - 27%

Abbreviations: FN, *false negative*; FP, *false positive*; TC, *trauma center*; TN, *true negative*; TP, *true positive*. *Accuracy metrics are presented as percentages and corresponding 95% Agresti-Coull binomial confidence intervals.

of digital decision-support and precision medicine, it is time to remove these abstractions and anticipate directly on the distribution and availability of resources. A competent reference standard is crucial in order to achieve this goal.

REFERENCE STANDARDS

The reference standard used to classify the patient's resource-need is the mainstay of trauma system evaluation. The most widely used reference standard is an Injury Severity Score (ISS) of 16 or greater.[6] The ISS, that builds upon the Abbreviated Injury Scale (AIS), was developed in 1974 to predict patient survival.[7] Apart from the limitations inherent in its formula (e.g., only the highest score per body region is taken into account), it comes as no surprise that it does not fully correlate with resource use.[8,9] Advantages of the ISS are its widespread implementation in trauma registries, capacity for retrospective analysis, and the known association between ISS, trauma center level, and mortality.[3]

A different approach to field triage is to assess a patient's resource-need at the scene of injury based on a set of predefined resources. Since the early 90s, several combinations of resources have emerged to be used as reference standards instead of the ISS.[8] Evaluation of the pre-hospital trauma system using resource-based reference standards removes the abstractions imposed by the ISS, but combining multiple resources into one endpoint does come at a price. It is then assumed that each resource is equally important to patient outcomes and is equally needed by the patient.[10] These assumptions cannot always be supported. Overall undertriage rates could be heavily influenced by a single resource (e.g., the overall undertriage rate in the Table is extremely dependent on Resource 1) and undertriage on one resource might be less important than other resources (e.g., damage control laparotomy versus blood transfusion).

Another major issue emerges when one aims to evaluate triage accuracy using these reference standards and available data. It is possible to construct a contingency table based on current resource use, but it is completely dependent on the current strategy. A patient in need of a specific resource may not have used it because it was unavailable (e.g., a patient was in need of a craniotomy but was erroneously transported to a facility without neurosurgical resources). The real question is: which resources should have been available considering the patient's injuries?

PERSONALIZED FIELD TRIAGE

Targeted improvement of field triage requires detailed insights into all distinct components of the triage strategy. Trauma systems should typically collect information on the following components: judgment of the patient's resource-need, decision-support of diagnostic tools, the supposed optimal hospital and reasons to bypass this hospital, where applicable.

Since trauma center-level designation criteria are based on available human and institutional resources, the obvious option would seemingly be to select the same resources for reference standards.[5] One possibility for a personalized reference standard, that builds upon existing infrastructure within trauma systems, is to classify whether a selected resource should be available (but not necessarily used) for the optimal treatment of each injury covered by the AIS. This process would be similar to the derivation of severity indicators for AIS codes (the post-dot part of the code ranging from one to six). Instead of severity indicators, a vector of resources must be derived in this case. Expert opinion could kickstart this process and evidence on the effect of resource availability on patient outcomes for a specific injury can be put in place at later points in time. Triage tools must be adapted to predict the selected resources rather than surrogate markers such as ISS or composite endpoints. The determination of the optimal transportation destination should then be based on the predicted resource-need (the professional's judgment and triage tool) and the availability of the specified resources within a trauma system. Accuracy metrics, such as undertriage and overtriage, can be evaluated for each selected resource and per patient. Overall undertriage, although less informative than resource-specific undertriage, is then defined as a patient transported to a hospital that lacks the resources that should have been available considering the patient's injuries.

This concept combines the advantages of existing anatomical and resource-based reference standards, gains insights in the accuracy and weaknesses of each component of the triage strategy and deals with major issues that hinder current triage evaluation. Once these steps have been implemented, they could provide a solid basis for a dynamic, learning, and personalized approach to field triage.

REFERENCES

1. Nathens AB, Jurkovich GJ, Maier RV, et al. **Relationship between trauma center volume and outcomes.** JAMA 2001; 285(9): 1164-71.

2. 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(1): 173-8.

3. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

4. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

5. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Aug 20, 2019).

6. Baker SP, O'Neill B. **The injury severity score: an update.** J Trauma 1976; 16(11): 882-5.

7. Gennarelli T, Wodzin E. **The Abbreviated Injury Scale 2005: update 2008**. Association for the Advancement of Automotive Medicine; Des Plaines, IL: 2008.

8. Baxt WG, Jones G, Fortlage D. **The trauma triage rule: a new, resource-based approach to the prehospital identification of major trauma victims.** Ann Emerg Med 1990; 19(12): 1401-6.

9. Newgard CD, Hedges JR, Diggs B, Mullins RJ. **Establishing the need for trauma center care: anatomic injury or resource use?** Prehosp Emerg Care 2008; 12(4): 451-8.

10. Tomlinson G, Detsky AS. **Composite end points in randomized trials: there is no free lunch.** JAMA 2010; 303(3): 267-8.

CHAPTER XI

General discussion

Table The 15 study questions and answers in thesis

Chapter	Study question and answer
2	What is the accuracy of pre-hospital trauma triage based on the initial transportation destination in a pediatric population around the world?
	The accuracy of pre-hospital trauma triage based on the initial transportation destination in children was unknown.
2	Which field triage tools that aid Emergency Medical Services professionals in the determination of the initial transportation destination were externally validated in a pediatric population and what is their diagnostic accuracy?
	Three different tools (Pediatric Trauma Triage Checklist, Trauma Scorecard, Field Triage Decision Scheme) were externally validated in children and their accuracy was poor to moderate at best.
3	What is the accuracy of pre-hospital trauma triage based on the initial transportation destination in a pediatric population in the Netherlands?
	The accuracy of pre-hospital trauma triage based on the initial transportation destination in children is insufficient.
3	How do contemporary field triage tools perform in terms of accuracy in a pediatric population based on an anatomical and a resource-based reference standard?
	The accuracy of contemporary field triage tools is too low to identify children in need of specialized trauma care.
4	What is the accuracy of pre-hospital trauma triage based on the initial transportation destination in an adult population in the Netherlands?
	The accuracy of pre-hospital trauma triage based on the initial transportation destination in adults is insufficient.
4	What is the accuracy of the National Protocol of Ambulance Services to select adults in need of specialized trauma in the pre-hospital setting based on an anatomical reference standard?
	Approximately two-thirds of adults in need of specialized trauma was not identified by the National Protocol of Ambulance Services.
5	How can we establish a prospective, open, and voluminous cohort to allow continuous monitoring of the pre-hospital trauma system?
	Patient suspected of injuries need to be identified from pre-hospital electronic health records. Included records need to be linked to in-hospital data collected by the Dutch National Trauma Registry.
5	What is the performance of a prediction model developed to select trauma patients from unfiltered pre-hospital electronic health records?
	An ensemble machine learning model (a combination of a Recurrent Neural Network and Logistic Regression) can very accurately discriminate between trauma patients and non-trauma patients.
5	What is the performance of a prediction model developed to link pre-hospital electronic health records to in-hospital patient outcomes collected by the Dutch National Trauma Registry?
	A Gradient Boosting Decision Tree trained to link electronic health records to in-hospital patient records can very accurately discriminate between matching records and non-matching records.
6	What is the predictive performance and external validity of a model to select severely injured patients in the pre-hospital setting based on an anatomical reference standard?
	The discriminative ability of a newly developed model was good and predicted probabilities are sufficiently calibrated after an intercept update.
7	What would be an adequate strategy to develop and validate a prediction model to identify patients in need of specialized trauma care?

	A combination of internal-external cross-validation and individual participant data meta-analysis can be used to develop transportable prediction models.
8	What is the predictive ability and external validity of novel prediction models to predict injury severity and patients' resource use?
	The model performance of two novel Gradient Boosting Decision Trees to predict injury severity and early critical-resource use is excellent. A simple calibration-in-the-large model is proposed to adapt the models to local settings.
8	What is the net benefit of the new prediction models compared to contemporary field triage tools?
	The net benefit of the newly developed models is greater or equal on the full range of threshold probabilities.
9	How could the impact of the Trauma Triage App be assessed in daily practice?
	The impact of the Trauma Triage App can be assessed in a multi-site stepped-wedge cluster randomized trial.
10	How can we optimize and evaluate field triage in inclusive trauma systems in the future?
	Evaluation of the pre-hospital trauma systems should be based on a patient's resource-need and the initial transportation destination. We argue that each resource needs to be evaluated independently.

CONCLUSIONS, IMPLICATIONS FOR PRACTICE, AND FUTURE PERSPECTIVES

This thesis provides answers to several important questions on the pre-hospital trauma triage strategy and the evaluation of pre-hospital trauma systems (Table). The following sections provide an overview of the conclusions, implications of our findings, and future perspectives.

PART I: EVALUATION OF FIELD TRIAGE

We showed that there was little research on field triage of injured children. No studies investigating destination-based accuracy rates were found in the literature. Consequently, it remained unknown whether children were transported to the right hospitals or not. In addition, we showed that none of the identified triage tools was accurate in children. The Pediatric Pre-hospital Trauma Triage (P2-T2) study eventually demonstrated that too many children in need of specialized care were erroneously transported to lower-level or non-pediatric trauma centers. We found that the lack of triage accuracy was even more striking in an adult population, although this was solely based on an anatomical reference standard in a more select population.

Our evaluation of the National Protocol of Ambulance Services (Choice of Hospital version 8.1; in Dutch, *Landelijk Protocol Ambulancezorg*) demonstrated that this actively used tool lacked the sensitivity to identify patients in need of specialized trauma care for both children and adults. The Field Triage Decision Scheme (version of 2011) of the American College of Surgeons Committee on Trauma (ACSCOT) turned out to be a slightly better alternative in children but was nevertheless insufficiently accurate.

Findings in this part of this thesis demonstrate a need for improved efforts to get the right patient to the right hospital. Contemporary triage tools lack accuracy and need to be replaced with – preferably personalized – alternatives. Other important factors that might lead to mistriage (e.g., trauma center proximity and patient acuity) are currently evaluated in the Trauma Triage Using Supervised Learning Algorithms (TESLA) trial. Efforts should be made to educate EMS professionals on the importance of field triage and the consequences of mistriage. We cannot stress enough that EMS professionals need to receive feedback on their triage decisions. Errors in pre-hospital patient management or decision-making need to be backpropagated to the responsible caregivers in order to improve the pre-hospital trauma system.

We argue that destination-based triage accuracy should be evaluated for each EMS region independently, in contrast to the prevailing belief that mistriage rates need to be computed for entire inclusive trauma systems or trauma regions. Undertriage and overtriage rates based on all patients admitted to trauma-receiving emergency departments are biased by self-transported and non-referred patients. Moreover, multiple EMSs might serve a single inclusive trauma region and aggregated mistriage rates will not generalize to individual EMSs. Guidelines for trauma system evaluation, such as the *Resources for Optimal Care of the Injured Patient* of the ACSCOT and the 2015 report of the Dutch Healthcare Institute (in Dutch, *Zorginstituut Nederland*), entitled *Spoed moet goed – indicatoren en normen voor zes spoedzorgindicaties*, should set norms for EMSs instead of, or in addition to, standards for trauma systems. After all, the decision of the initial transportation destination is made by EMS professionals.

PART II: THE TRAUMA CONTINUUM OF CARE COHORT

This part described the methodology used to construct the Trauma Continuum of Care Cohort (TRACCC). Two tools that were invaluable to the research presented in this thesis were *SelectAssist* and *LinkAssist*. SelectAssist proved to be a highly accurate prediction model to select trauma patients from unfiltered electronic health records. LinkAssist, on the other hand, was also very accurate and may be used to link pre-hospital electronic

health records to data from the Dutch National Trauma Registry (in Dutch, *Landelijke Traumaregistratie [LTR]*). The combination of these tools enabled us to construct a highly voluminous cohort based on 1.5 million pre-hospital patient records. Trauma registrars, in contrast, manually link pre-hospital records to in-hospital data to collect pre-hospital vital signs. LinkAssist was built to automate or aid this tedious process and can save huge amounts of time and costs. Efforts are made to improve both SelectAssist and LinkAssist. We plan to implement categorical classification of medical specialties (e.g., neurology, cardiology) in SelectAssist to facilitate its use in non-trauma research. Moreover, we suspect that its model's performance might be improved with the adoption of transfer learning techniques in natural language processing, such as BERT or ELMo.[1,2]

Limitations of TRACCC are inherent to its original data sources. First, missing data are empirical in time-critical situations such as field triage. Second, the LTR only covers the first early critical-resource that was used. Third, the LTR does not include patients that die on the scene of injury or during transportation. We illustrate the importance of these data in *Part III* and hope that future revisions of the data template of the LTR will reconsider these design decisions.

TRACCC is arguably the most important outcome of this thesis. The potential of TRACCC has only been minimally explored and exploited in the context of this thesis. This cohort enables researchers to study the pre-hospital trauma system on an unprecedented scale. For instance, researchers might use TRACCC to (1) study the effect of trauma center care or pre-hospital care on patient outcomes, (2) conduct geospatial analyses (e.g., effect of trauma proximity on patient outcomes), and to (3) evaluate the three phases of triage in inclusive trauma systems. We hope that more regions will participate in data collection for TRACCC and that it will eventually become integral to the LTR to empower continuous monitoring of the pre-hospital trauma system in the Netherlands.

PART III: PERSONALIZED FIELD TRIAGE

We developed and validated a prediction model to identify severely injured patients at the scene of injury. Model performance was good, and predicted probabilities were well-calibrated after the application of a simple model updating technique. I developed a smartphone and tablet application that incorporated this prediction model to facilitate its use at the scene of injury (i.e., the Trauma Triage App; Figure 1). The impact of the Trauma Triage App on triage accuracy is currently evaluated in eight EMS regions in the Netherlands (TESLA trial).

Limitations of the Trauma Triage App were addressed in the Gradient Boosted Trauma Triage (GOAT) study. Two prediction models to (i) predict injury severity and (ii) early critical-resource use were presented in this study. Both models outperformed contemporary triage tools in terms of model performance. These models do not require complete data availability and are able to dynamically update predictions whenever new data become available.

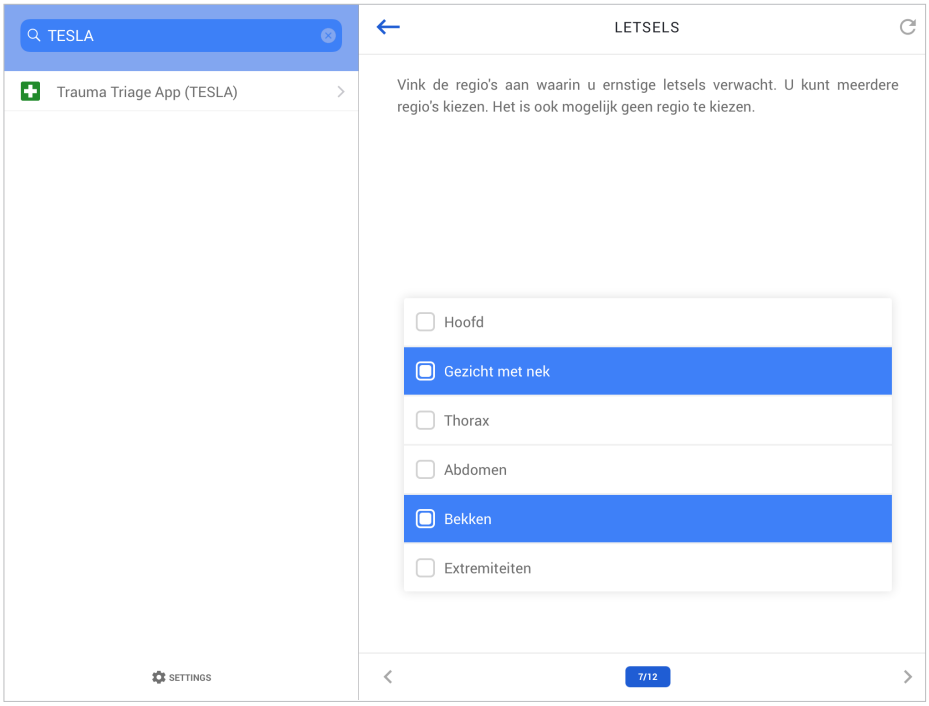


Figure 1 User Interface of the Trauma Triage App (iPadOS)

The Trauma Triage App was our first effort to map and improve the pre-hospital triage strategy in the Netherlands. This smartphone application requires EMS professionals to fill out questions on their judgment, preferred transportation destination, and if applicable, reasons to bypass this destination (e.g., distance). Moreover, it was the first effort to provide EMS professionals with personalized predictions to aid the identification of severely injured patients (Figure 2). This signified an important change in field triage. More than two centuries after Percy supposedly established the first triage paradigm, we developed and implemented a non-categorical triage strategy.[3] Models constructed in the GOAT study symbolize the next generation of triage tools: hands-free, integrated

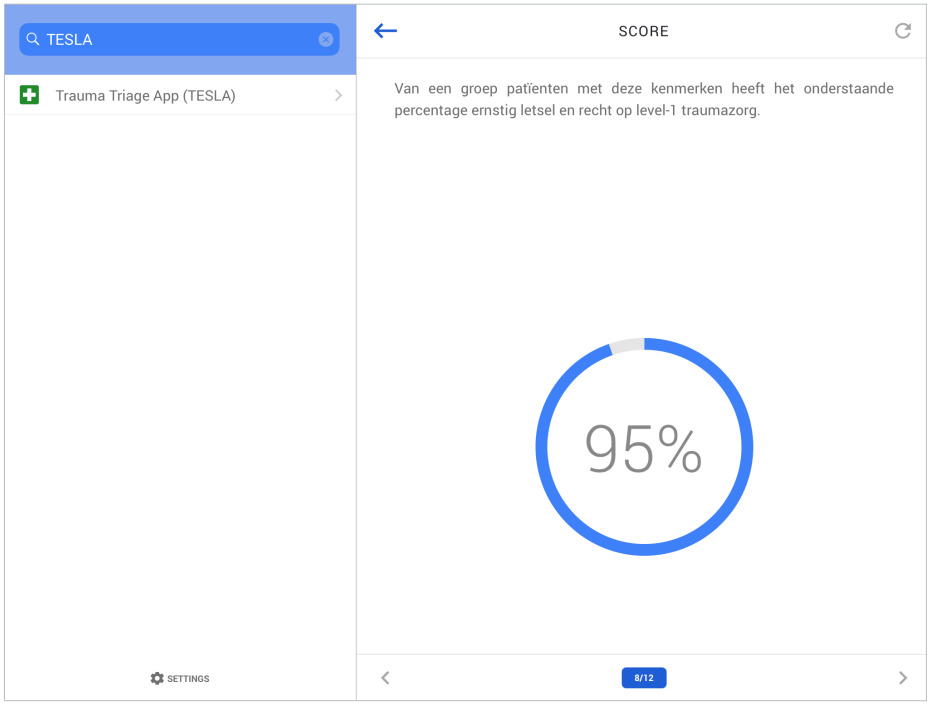


Figure 2 Personalized Predictions Computed by the Prediction Model of the Trauma Triage App

with pre-hospital electronic health records, and preferably used to predict resource use. These models were designed to work in adverse conditions with limited time, few diagnostic modalities, and incomplete data. We hope that they will succeed to get the right patient to the right hospital, either as a primary triage tool or integral to Early Warning Systems.

PART IV: DYNAMIC TRAUMA SYSTEMS

The final part of this thesis described the current dichotomous view on triage in modern trauma systems and proposed new methodology to monitor and optimize the pre-hospital trauma system. Traditional triage tools classified patients as high (severely injured) or low-risk (non-severely injured) and classified trauma centers as higher or lower-level. We argued that reference standards need to reflect a patient's resource-need (instead of resource use) and uncovered several limitations of the Injury Severity Score in the context of field triage evaluation. Finally, we proposed a new concept to evaluate the accuracy of pre-hospital trauma systems. In our opinion, we need to evaluate triage accuracy per

resource for each patient individually. The determination of the optimal facility should be based on the predicted resource-need at the scene of injury and the dynamic availability of resources within a trauma system (Figure 3). This would be fundamental to a dynamic supply and demand model, that might be able to unlock the true potential of inclusive trauma systems.

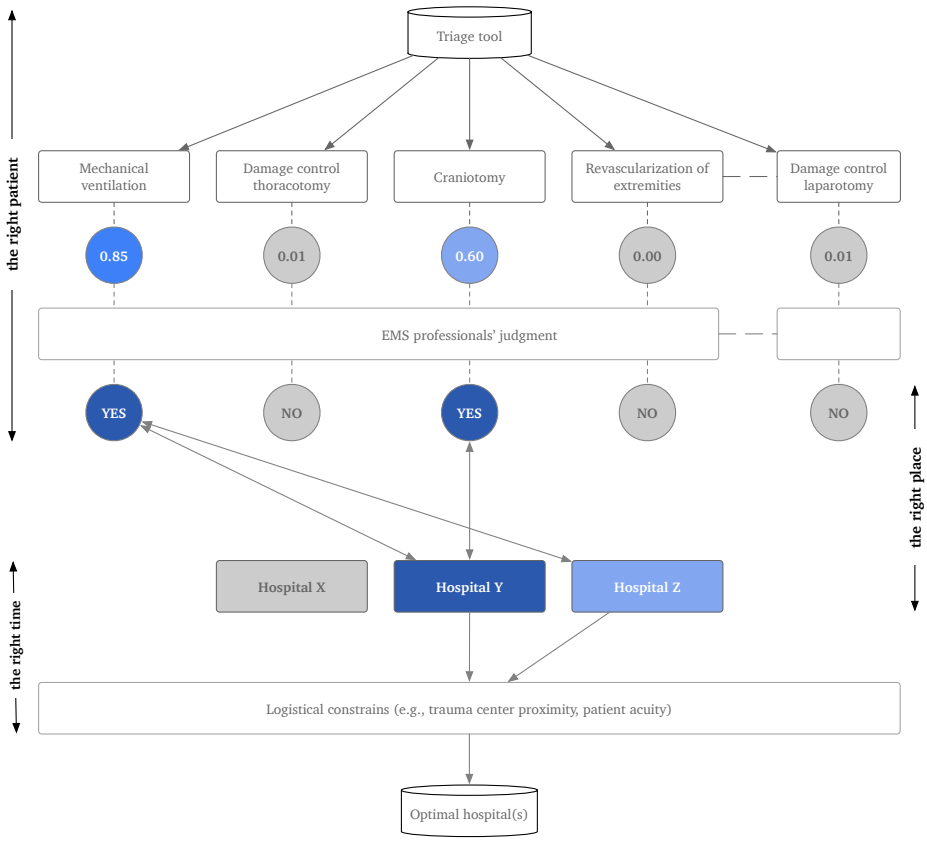


Figure 3 A Theoretical Model for a Resource-Based Approach to Pre-Hospital Trauma Triage. First, triage tools generate predictions for a set of pre-specified resources. Second, EMS professionals adhere to, or overrule the triage tool's advice. Third, surrounding hospitals with sufficient resources and expertise are selected. Finally, the initial transportation destination is determined in light of step 1-3 and logistical constraints.

FUTURE PERSPECTIVES AND RECOMMENDATIONS

In 2017, all 25 EMSs employed a combined total of 2284 paramedics.[4] Approximately 4500 patients were considered to be severely injured (ISS ≥16) that year.[5] This implies that EMS professionals normally encounter less than two severely injured patients every year. This observation, current undertriage rates, and an opaque triage strategy emphasize the need for change:

1. *The right patient.* Pre-hospital trauma triage should focus on the identification of patients that potentially need to use specialized resources (e.g., damage control surgery). We envision that triage tools need to be able to predict resource-need based on routinely collected data. These tools should be tightly integrated with electronic health records, highly accurate, and user-friendly. We argue that EMS professionals need to decide upon the final set of potentially needed resources in light of all available information (including triage tools).
2. *The right place.* Up-to-the-minute information on resource availability should be available within the trauma system to determine the optimal transportation destination. A patient's resource-need can then be matched with hospitals with sufficient resources and expertise for optimal treatment.
3. *The right time*. Global positioning systems can be used to determine the expected travelling distance and time at the scene of injury to surrounding hospitals (Figure 4). Optimal hospitals, travelling times, and patient acuity should be balanced by EMS professionals to determine the initial transportation destination.
4. *Transparence*. Pre-hospital decision-making should be transparentized. Reasons to choose a hospital, overrule a triage tool, or bypass a hospital should be documented in the electronic health record.
5. *Evaluation.* Mistriage rates should be calculated for each resource and each EMS region separately. Norms should be set for EMSs and emergency departments need to provide EMS professionals with feedback on pre-hospital decision-making in general and field triage in particular.

This thesis presented clinical research on the evaluation and optimization of pre-hospital trauma systems that was conducted between 2015 and 2020. Important discoveries were made, new tools were developed, and new concepts were proposed. Yet, much work remains to be done.

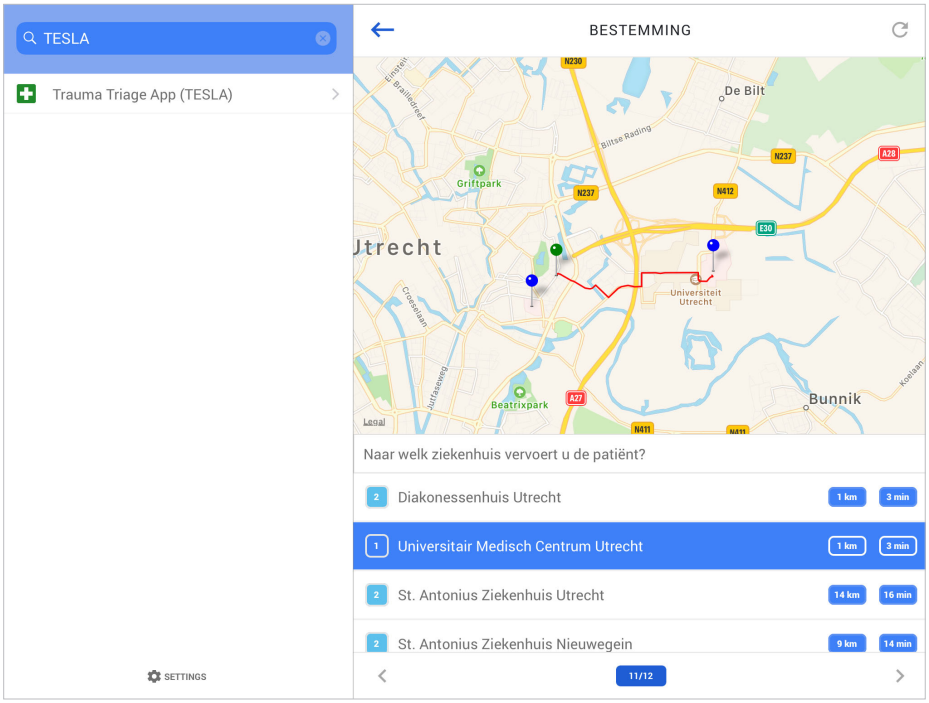


Figure 4 Travelling Time and Distance to Surrounding Hospitals

REFERENCES

1. Devlin J, Chang MW, Lee K, Toutanova K. BERT: **Pre-training of deep bidirectional transformers for language understanding.** arXiv:181004805 2018.

2. Peters ME, Neumann M, Iyyer M, Gardner M, Clark C. **Deep contextualized word representations.** Proceedings of NAACL 2018.

3. Nakao H, Ukai I, Kotani J. **A review of the history of the origin of triage from a disaster medicine perspective.** Acute Med Surg 2017; 4(4): 379-84.

4. Ambulancezorg Nederland. **Tabellenboek 2017**. Zwolle, 2018.

5. Landelijke Netwerk Acute Zorg. **Landelijke Traumaregistratie 2013 - 2017.** Utrecht, 2018.

CHAPTER XII

Summary

Injury remains a leading cause of death and disability on a global scale.[1] The treatment and organization of care of severely injured patients has changed considerably over the last decades.[2] Lessons learned during the French Revolution and in the Napoléonic Wars eventually led to the development of inclusive trauma systems.[3] These regionalized efforts to centralize resources and expertise proved to decrease mortality rates and injury-related readmissions.[4-7]

The greater the amount of centralization within inclusive trauma systems, the greater the importance of adequate field triage. The ultimate goal of field triage is to get the right patient to the right place in the right time. Transporting patients in need of specialized care to lower-level trauma centers (i.e., undertriage) is associated with increased mortality rates, whereas overtriage – the transportation of patients not requiring specialized trauma care to higher-level trauma centers – is not considered to be cost-effective.[4-8] An impeccable pre-hospital trauma system should hypothetically be able to eliminate unintentional mistriage.

The Netherlands implemented a nationwide inclusive trauma system in the late 1990s.[9] The Dutch trauma system consists of 11 inclusive trauma regions that are each coordinated by a single level-I (i.e., higher-level) trauma center. Every inclusive trauma region encompasses multiple level-II/III trauma centers (i.e., lower-level).

The prevalence of patients in need of specialized trauma care is generally low. Single region studies are often underpowered to study diagnostic strategies, causal treatment effects, and the impact of the organization of trauma care on patient outcomes. Institutional and regional collaboration is key to study pre-hospital trauma systems. In 2019, we established the Pre-hospital Trauma Triage Research Collaborative (PTTRC) with eight Emergency Medical Services (EMS) regions and seven inclusive trauma regions to study effects of triage that transcend the pre-hospital setting. The PTTRC is an early initiative to cooperate in research on the Dutch pre-hospital trauma system and organization of trauma care.

PART I: EVALUATION OF THE PRE-HOSPITAL TRAUMA TRIAGE STRATEGY

Advancement of protective measures, diagnosis, treatment, and organization of care of severely injured patients reduced mortality rates in the last decades but field triage is mainly unexplored territory. Yet, field triage is considered to be a crucial step in the chain of critical trauma care because all the other steps depend on it.

The relative burden of injuries is highest in children.[10] Unintentional injuries are the greatest cause of death in children over five years and injury-related disabilities can have long-lasting impact on learning, relationships, and life in general.[10] In *Chapter II,* we evaluated the accuracy of pediatric triage strategies around the world by means of a systematic review of the literature. Only five studies with poor to mediocre methodological quality were identified with a combined total of 1222 severely injured children. No study investigated triage rates based on the initial transportation destination and none of the triage tools was both highly sensitive (>95%) and specific (>65 – 75%).

In *Chapter III* we conducted an observational study to investigate the accuracy of pediatric trauma triage based on the initial transportation destination. In addition, we explored the external validity of two triage tools used in contemporary practice: the Dutch National Protocol of Ambulance Services (Choice of Hospital, version 8.1) and the Field Triage Decision Scheme (version of 2011) of the American College of Surgeons Committee on Trauma (ACSCOT).[2,11] We found that one of six-to-seven patients in need of specialized trauma care was not transported to a level-I pediatric trauma center. These triage rates do not comply with the guidelines of Dutch Healthcare Institute that demand undertriage rates of less than 10%.[12] The triage tools evaluated in this study were inaccurate in the Dutch population with sensitivity rates of maximally 64%. This implies that at least one of three children in need of specialized care is missed by both triage tools.

A similar study in an adult population was presented in *Chapter IV*. This study only included patients that were transported within a single EMS region with the highest priority available and demonstrated an undertriage rate of over 20%. This study evaluated the Dutch triage protocol identically to *Chapter III* and reported that the protocol did not recognize approximately two of three severely injured patients.

PART II: THE TRAUMA CONTINUUM OF CARE COHORT

Accurate and effective evaluation of field triage requires data from a consecutive or random sample of patients transported by EMSs. Trauma patients, injury patterns, and EMS regions can be extremely heterogeneous. In *Chapter V*, we present the Trauma Continuum of Care Cohort (TRACCC) that set out to study pre-hospital field triage and decision-making in multiple regions in the Netherlands. Two tools were developed to aid the construction of the cohort: *SelectAssist* is an ensemble machine learning model that was used to select trauma patients from EMS electronic health records and *LinkAssist* was used to provide computationally feasible and highly accurate linkage of pre-hospital and in-hospital patient records.

TRACCC includes data from eight EMSs, seven inclusive trauma systems, and 65 trauma-receiving emergency departments. A total of 165,404 injured patients – of which approximately 3800 required specialized trauma care – were included between Jan 2015 and Dec 2017. TRACCC covers data on demographic characteristics, vital signs, mechanism of injury, and in-hospital patient outcomes. The goal of TRACCC was to continuously monitor, research, and optimize field triage on a national scale although its scope allows researchers to study the entire pre-hospital trauma system.

PART III: PERSONALIZED FIELD TRIAGE

In the first part of this thesis we uncovered the flaws in contemporary triage tools. Our first effort to create a new triage tool is presented in *Chapter VI*. We selected pre-defined predictors to construct a new prediction model using penalized maximum likelihood estimation. This model was based on an anatomical reference standard, individual participants data (IPD) from a single EMS region and inclusive trauma region, and was externally validated on IPD from a different EMS region. This prediction model was able to attain a sensitivity of 89% and a specificity of 50%[3-7] depending on the threshold value.

In *Chapter VII* and *VIII*, we present the design and the results of the Gradient Boosted Trauma Triage (GOAT) study. The GOAT study aimed to develop and validate prediction models to identify patients in need of specialized care during field triage based on automated measurement and routinely collected data. We adopted an internal-external cross-validation strategy to construct two new gradient-boosting decision trees: (i) one prediction model based on the Injury Severity Score of 16 or greater, and (ii) a second model based on early critical-resource use. An IPD meta-analysis was performed to investigate heterogeneity in model performance (e.g., concordance statistics, observed

versus expected ratios, calibration slopes). The resulting models were highly accurate, superior to the model constructed in *Chapter VI* and contemporary triage tools, and were able to deal with the presence of missing data.

In *Chapter IX*, we outline the design of stepped wedge cluster randomized trial (TESLA trial) to evaluate the impact of the Trauma Triage App on triage accuracy in pre-clinical practice. I developed the Trauma Triage App (version 2) that incorporated the prediction model presented in *Chapter VI* and a questionnaire to make the triage strategy more transparent. The TESLA trial is currently ongoing in eight EMS regions and is expected to complete enrolment at the end of 2020.

PART IV: DYNAMIC TRAUMA SYSTEMS

The fallacies and flaws of present-day triage strategies in pre-hospital trauma systems are outlined in *Chapter X*. We argue that the Injury Severity Score should not be used to evaluate the accuracy of field triage. Instead, we suggest a resource-based approach to get the right patient to the right hospital. Finally, we present a new concept that builds upon existing anatomical and resource-based reference standards. In our opinion, triage tools should be used to predict potential resource-need for individual patients and triage accuracy should be evaluated for each resource separately. One way to achieve this would be to derive a set of potentially needed resources for each injury described in the Abbreviated Injury Scale manual.[13] The ultimate goal of this concept is to unlock the true potential of precision medicine in pre-hospital trauma systems.

REFERENCES

1. Haagsma JA, Graetz N, Bolliger I, et al. **The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013.** Inj Prev 2016; 22(1): 3-18.

2. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Dec 20, 2019).

3. Baker D, Cazalaa JB, Carli P. Resuscitation great. **Larrey and Percy--a tale of two barons.** Resuscitation 2005; 66(3): 259-62.

4. Celso B, Tepas J, Langland-Orban B, et al. **A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems.** J Trauma 2006; 60(2): 371-8; discussion 8.

5. Curtis K, Chong S, Mitchell R, Newcombe M, Black D, Langcake M. **Outcomes of severely injured adult trauma patients in an Australian health service: does trauma center level make a difference?** World J Surg 2011; 35(10): 2332-40.

6. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

7. Staudenmayer K, Weiser TG, Maggio PM, Spain DA, Hsia RY. **Trauma center care is associated with reduced readmissions after injury.** J Trauma Acute Care Surg 2016; 80(3): 412-6; discussion 6-8.

8. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

9. Twijnstra MJ, Moons KG, Simmermacher RK, Leenen LP. **Regional trauma system reduces mortality and changes admission rates: a before and after study.** Ann Surg 2010; 251(2): 339-43.

10. WHO. **World report on child injury prevention.** Geneva, 2008. https://www.who.int/violence_injury_prevention/child/injury/world_report/en/ (accessed Dec 20, 2019).

11. Ambulancezorg Nederland**. Landelijk Protocol Ambulancezorg (versie 8.1).** Zwolle, 2016. https://www.ambulancezorg.nl/themas/kwaliteit-van- zorg/protocollen-en-richtlijnen/landelijk-protocol-ambulancezorg (accessed Dec 20, 2019).

12. Zorginstituut Nederland. **Spoed moet goed: indicatoren en normen voor zes spoedzorgindicaties;** 2015.

13. Gennarelli T, Wodzin E. **The Abbreviated Injury Scale 2005: update 2008**. Association for the Advancement of Automotive Medicine; Des Plaines, IL: 2008.

CHAPTER XIII

Summary in Dutch

Samenvatting in het Nederlands

Ongevallen zijn een belangrijke oorzaak van overlijden en permanente invaliditeit.[1] De behandeling van letsels en de organisatie van trauma zorg is in de laatste decennia aanzienlijk veranderd.[2] Lessen die werden getrokken uit het behandelen en vervoeren van ongevalspatiënten in oorlogssituaties vormden de basis voor het ontstaan van geregionaliseerde trauma zorg, ofwel 'inclusieve trauma systemen'.[3] De inspanning om middelen en expertise inzake de behandeling van ernstig gewonde patiënten te concentreren heeft er uiteindelijk voor gezorgd dat de ongevalsterfte daalde en heeft evenzeer tot een reductie van het aantal heropnamen geleid.

Volwassen trauma systemen hebben veelal een grotere mate van centralisatie van middelen ten opzichte van recentere initiatieven om trauma zorg te regionaliseren. De mate van centralisatie is bepalend voor het belang van adequate triage van ongevalspatiënten op de plaats van het ongeval. Het uiteindelijke doel van deze vorm van triage (uit het Frans *trier* = 'sorteren') is om de juiste patiënt, binnen de juiste tijd, naar het juiste ziekenhuis te transporteren. Het vervoeren van patiënten met de behoefte aan gespecialiseerde trauma zorg naar een ziekenhuis zonder voldoende expertise en middelen voor optimale behandeling (ondertriage) is logischerwijs geassocieerd met nadelige gezondheidsuitkomsten voor patiënten, zoals een verhoogde kans op sterfte en blijvende letsels.[4-7] Tegenovergesteld wordt overtriage – het vervoeren van patiënten met mild en matige letsels naar gespecialiseerde traumacentra – in verband gebracht met een toename van kosten.[8] Hypothetisch zou een trauma systeem onopzettelijke foutieve triage kunnen reduceren tot nul. In de praktijk is het wellicht mogelijk om de ondertriage drastisch te beperken, maar geldt dit niet voor de overtriage in het huidige trauma systeem. De relatief lage prevalentie van ernstige letsels ten opzichte van het aantal gespecialiseerde (level-I) traumacentra weerhoudt deze centra ervan om zich te 'superspecialiseren'.

Twintig jaar geleden heeft er in Nederland regionalisatie van de trauma zorg plaatsgevonden.[9] Het land werd onverdeeld in 11 trauma regio's die elk gecoördineerd worden door een enkel level-I traumacentrum. Elk van deze 'inclusieve' trauma regio's biedt trauma zorg voor het volledige spectrum aan ongeval gerelateerde letsels. Om trauma systemen te onderzoeken is er een grote hoeveelheid, heterogene en hoogkwalitatieve data nodig afkomstig uit verschillende regio's. Interregionale en multidisciplinaire studies zijn de sleutel tot de verbetering van trauma zorg. Het doel van dit proefschrift was niet enkel om antwoorden te vinden op belangrijke vraagstukken uit de pre-hospitale trauma zorg, maar evengoed een poging om een draagvlak te creëren voor regio-overstijgend onderzoek naar het pre-hospitale trauma systeem.

DEEL I: EVALUATIE VAN DE PRE-HOSPITALE TRAUMA TRIAGE STRATEGIE

Ontwikkelingen op het gebied van ongevalspreventie, diagnostische modaliteiten, behandelstrategieën en de organisatie van zorg heeft ertoe geleid dat de ongevalssterfte in de laatste jaren gedaald is. De noodzaak om lasten ten gevolge van ongevalsproblematiek te reduceren blijft echter onverminderd hoog.

De impact van ongevallen op kinderen is relatief het grootst.[10] Letsels zijn de grootste oorzaak van kindersterfte (>5 jaar) en invaliditeit ten gevolge van een ongeval kan langdurige gevolgen hebben op het vermogen tot leren, het vormen van sociale contacten en op het leven in zijn algemeenheid.[10] In *Hoofdstuk II* hebben we de accuratesse van pre-hospitale triage van kinderen onderzocht door middel van een systematische literatuurstudie. Slechts vijf onderzoeken van wisselende kwaliteit konden worden geïdentificeerd. Geen enkele studie evalueerde de daadwerkelijke accuratesse van de volledige triage strategie op basis van het 'level' van het ziekenhuis waar de patiënt initieel naar vervoerd werd. Geen van de triageprotocollen die extern gevalideerd werden was zowel sensitief (>95%) als specifiek (>65 – 75%).

In *Hoofdstuk III* hebben we in een interregionale studie in Nederland onderzocht wat de nauwkeurigheid van de volledige pre-hospitale triage strategie voor kinderen was. Daarnaast hebben we de externe validiteit van twee actief gebruikte triage hulpmiddelen geëvalueerd: het Landelijk Protocol Ambulancezorg ('Keuze ziekenhuis', versie 8.1) en het Amerikaanse Field Triage Decision Scheme (versie uit 2011).[2,11] Onze resultaten toonden dat één op de zes à zeven patiënten die behoeftig waren aan gespecialiseerde trauma zorg niet getransporteerd werd naar een geschikt (level-I) traumacentrum voor kinderen. Het Zorginstituut Nederland publiceerde eind 2015 een richtlijn met een norm voor ondertriage, te weten minder dan 10%.[12] De resultaten uit deze studie demonstreerden dat we nog niet in staat zijn om deze richtlijn na te leven. De triage hulpmiddelen die getest werden in dit onderzoek waren maximaal in 64% van de gevallen sensitief. Dit impliceert dat meer dan één op de drie ernstig gewonde patiënten niet als dusdanig herkend wordt door de huidige protocollen.

Een soortgelijk onderzoek hebben we uitgevoerd in een populatie met volwassen trauma patiënten. Deze studie, gepresenteerd in *Hoofdstuk IV*, includeerde enkel patiënten die met de hoogst mogelijke prioriteit (sirenes en zwaailichten) naar een spoedeisende hulp werden getransporteerd door één ambulancedienst. De ondertriage die werd gevonden in deze studie bedroeg meer dan 20% en was daarmee meer dan tweemaal zo hoog als de maximale bovengrens van de gestelde norm. Overeenkomstig met *Hoofdstuk III* bleek

het Nederlandse triage protocol ook niet in staat te zijn om een goed onderscheid te maken tussen milde en ernstig gewonde patiënten: twee op de drie patiënten met ernstige letsels werd niet als dusdanig herkend.

DEEL II: HET TRAUMA CONTINUUM OF CARE COHORT

Minutieus onderzoek naar pre-hospitale triage vereist data van een volledige populatie of willekeurige steekproef van patiënten die vervoerd zijn door ambulancediensten. Trauma patiënten, letselpatronen, en ambulancediensten verschillen onderling substantieel van elkaar. Generaliseerbaar onderzoek vereist dan ook multiregionale data. In *Hoofdstuk V* presenteren we het Trauma Continuum of Care Cohort (TRACCC) dat speciaal ontworpen werd om pre-hospitale triage en beslisvorming te onderzoeken in verscheidene regio's in Nederland. Twee softwarematige hulpmiddelen werden ontwikkeld om het cohort te creëren: (i) *SelectAssist* (een 'ensemble machine learning' voorspelmodel) werd gebruikt om trauma patiënten te identificeren op basis van vrije teksten in elektronische patiëntendossiers, en (ii) *LinkAssist* werd gebruikt om pre-hospitale patiëntendossiers en ziekenhuisuitkomsten van patiënten aan elkaar te koppelen. Beide hulpmiddelen bleken zeer accuraat (concordantie statistiek van 1.0) in een ongeziene dataset.

TRACCC herbergt demografische karakteristieken, vitale parameters en ziekenhuisuitkomsten van 165,504 patiënten die vervoerd werden door acht ambulancediensten naar 65 verschillende ziekenhuizen in zeven van de 11 trauma regio's. Ongeveer 3800 van deze patiënten waren behoeftig aan gespecialiseerde trauma zorg. Het initiële doel van TRACCC was om de nauwkeurigheid van pre-hospitale triage continue te kunnen monitoren, te onderzoeken en te optimaliseren op nationaal niveau. De grootte en reikwijdte van TRACCC biedt echter de mogelijkheid om onderzoek te doen naar het gehele pre-hospitale trauma systeem.

DEEL III: GEPERSONALISEERDE PRE-HOSPITALE TRIAGE

In het eerste deel van dit proefschrift legden we de onvolkomenheden van huidige triage hulpmiddelen bloot. Onze eerste poging om een nieuw en nauwkeurig hulpmiddel te creëren werd gepresenteerd in *Hoofdstuk VI*. Voor dit hulpmiddel (voorspelmodel) selecteerden we voorspellende factoren op basis van klinisch redeneren en gebruikte we *penalized maximum likelihood* om model parameters te schatten. Dit voorspelmodel was

gebaseerd op data van individuele patiënten (IPD) afkomstig uit één trauma regio. Het ontwikkelde model werd extern gevalideerd op basis van IPD uit Brabant en was, afhankelijk van de drempelwaarde, sensitief in 89% van de ernstig gewonde patiënten waarbij de specificiteit 50% was.

In *Hoofdstuk VII* en *VIII* presenteerden we het ontwerp en de resultaten van de Gradient Boosted Trauma Triage (GOAT) studie. Deze studie werd opgezet om nieuwe voorspelmodellen te ontwikkelen met als doel de juiste patiënt naar het juiste ziekenhuis te transporteren. Hoewel dit doel overeenkomstig was met het predictiemodel dat reeds gepresenteerd werd in *Hoofdstuk VI* was de methodologie om dit te bewerkstelligen zeer verschillend. In de GOAT-studie maakten we gebruik van een strategie die interne-externe kruisvalidatie wordt genoemd en werd de kwaliteit van de ontwikkelde modellen (discriminatie en calibratie) vastgesteld door middel van een meta-analyse. Op basis van geautomatiseerde meetgegevens en routinematig verzamelde data uit TRACCC werden twee gradient-boosting decision trees ontwikkeld: (i) één voorspelmodel gebaseerd op de Injury Severity Score groter of gelijk aan 16 en (ii) een tweede model waarin het gebruik van gespecialiseerde ziekenhuiszorg centraal stond. De resulterende modellen waren zeer accuraat en een decision curve analyse toonde een verbetering ten opzichte van het model dat gepresenteerd werd in *Hoofdstuk VI* en het Landelijk Protocol Ambulancezorg. De modellen uit deze studie onderscheiden zich echter op een ander belangrijk vlak: zelfs zonder volledige informatie en zonder menselijke interactie kunnen deze modellen persoonlijke voorspellingen genereren. Dit is een essentiële kwaliteit om de integratie van een dergelijk model in een acute setting te bespoedigen.

In *Hoofdstuk IX* presenteren we het ontwerp van de Trauma Triage using Supervised Learning Algorithms (TESLA) studie. Deze studie beoogt de impact van het model dat gepresenteerd werd in *Hoofdstuk VI* te evalueren in de praktijk. Om dit model te kunnen testen heb ik de Trauma Triage App (versie 2) ontwikkeld. Deze smartphone applicatie voor Android en iOS stelt enkele vragen om de kans op ernstig letsel te schatten en poogt daarnaast om door middel van gerichte vragen in zicht te geven in de triage strategie. De TESLA-trial wordt momenteel uitgevoerd in acht van de 25 ambulanceregio's in Nederland en de verwachting is dat de resultaten begin 2021 zullen volgen.

DEEL IV: DYNAMISCHE TRAUMA SYSTEMEN

De mankementen en fundamentele problemen van hedendaagse triage strategieën werden besproken in *Hoofdstuk X*. In dit hoofdstuk stellen we dat de Injury Severity Score niet meer gebruikt zou moeten worden voor het evalueren van de accuratesse van pre-

hospitale triage. Als alternatief suggereren we een referentiestandaard die gebaseerd is op de potentiële behoefte aan gespecialiseerde middelen. Ten behoeve hiervan presenteerden we een nieuw concept dat gebaseerd is op zowel een anatomische als een op middelengebruik gebaseerde referentiestandaard. Triage hulpmiddelen zouden in onze ogen gebruikt moeten worden om de potentiële behoefte aan individuele middelen te voorspellen en op basis van deze behoefte dient er een ziekenhuis geselecteerd te worden met sufficiënte middelen en expertise. De nauwkeurigheid van triage dient vervolgens geëvalueerd te worden per middel in plaats van op basis van een samengevoegde referentiestandaard. Het uiteindelijke doel van dit concept is om het ware potentieel van precisie-geneeskunde in de pre-hospitale setting optimaal te benutten.

REFERENTIES

1. Haagsma JA, Graetz N, Bolliger I, et al. **The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013.** Inj Prev 2016; 22(1): 3-18.

2. American College of Surgeons Committee on Trauma. **Resources for optimal care of the injured patient: 2014.** Chicago, IL, 2014. https://www.facs.org/quality-programs/trauma/tqp/center- programs/vrc/resources (accessed Dec 20, 2019).

3. Baker D, Cazalaa JB, Carli P. Resuscitation great. **Larrey and Percy--a tale of two barons.** Resuscitation 2005; 66(3): 259-62.

4. Celso B, Tepas J, Langland-Orban B, et al. **A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems.** J Trauma 2006; 60(2): 371-8; discussion 8.

5. Curtis K, Chong S, Mitchell R, Newcombe M, Black D, Langcake M. **Outcomes of severely injured adult trauma patients in an Australian health service: does trauma center level make a difference?** World J Surg 2011; 35(10): 2332-40.

6. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. **A national evaluation of the effect of trauma-center care on mortality.** N Engl J Med 2006; 354(4): 366-78.

7. Staudenmayer K, Weiser TG, Maggio PM, Spain DA, Hsia RY. **Trauma center care is associated with reduced readmissions after injury.** J Trauma Acute Care Surg 2016; 80(3): 412-6; discussion 6-8.

8. Newgard CD, Staudenmayer K, Hsia RY, et al. **The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers.** Health Aff (Millwood) 2013; 32(9): 1591-9.

9. Twijnstra MJ, Moons KG, Simmermacher RK, Leenen LP. **Regional trauma system reduces mortality and changes admission rates: a before and after study.** Ann Surg 2010; 251(2): 339-43.

10. WHO. **World report on child injury prevention.** Geneva, 2008. https://www.who.int/violence_injury_prevention/child/injury/world_report/en/ (accessed Dec 20, 2019).

11. Ambulancezorg Nederland**. Landelijk Protocol Ambulancezorg (versie 8.1).** Zwolle, 2016. https://www.ambulancezorg.nl/themas/kwaliteit-van- zorg/protocollen-en-richtlijnen/landelijk-protocol-ambulancezorg (accessed Dec 20, 2019).

12. Zorginstituut Nederland. **Spoed moet goed: indicatoren en normen voor zes spoedzorgindicaties;** 2015.

PART X

Review committee

Acknowledgements (dankwoord)

Curriculum vitae

REVIEW COMMITTEE

Prof. dr. J. N. Brinkkemper
Professor of Organization and Information, Utrecht University

Prof. dr. M. J. R. Edwards
Professor of Trauma Surgery, Radboud University Medical Center

Prof. dr. ir. M. J. C. Eijkemans
Professor of Biostatistics, University Medical Center Utrecht

Prof. dr. F. C. Öner
Professor of Spinal Surgery, University Medical Center Utrecht

Prof. dr. M. R. Vriens
Professor of Endocrine Oncologic Surgery, University Medical Center Utrecht

CURRICULUM VITAE

Rogier van der Sluijs was born on July 27th, 1991 in Utrecht, the Netherlands. He was raised together with one sister in a supportive family in Odijk and graduated from the Gymnasium K.S.G. de Breul in Zeist in 2009. He spent the next year as a snowboard and ski-instructor in Saalbach-Hinterglemm, Austria.

In 2010 he started medical school at the University of Utrecht. He joined the newfound triage research group of Dr. M. van Heijl and Prof. Dr. L.P.H. Leenen at the Department of Surgery of the University Medical Center in Utrecht as a research assistant in 2015. Thereafter, he contributed to the design of the nationwide TESLA trial and initiated the Dutch Pre-hospital Trauma Triage Research Collaborative with his supervisory team. In September 2016, he was admitted to the MSc Epidemiology at the Graduate School of Life Sciences, Utrecht University. During this period, he conducted research projects on low and high-dimensional Machine Learning. He graduated from medical school in 2017 and contiguously continued his research as a Ph.D. student at the University Medical Center Utrecht and held a research position at the Maastricht University Medical Center+ under supervision of Prof. Dr. M. Poeze. In this timespan, he developed the Trauma Triage App, coordinated the ZonMw-funded TESLA trial, and constructed the Trauma Continuum of Care Cohort. He graduated from the MSc Epidemiology in 2019.

Rogier van der Sluijs started new projects on artificial intelligence in medicine in 2019 at the University Medical Center in Utrecht and will continue his research endeavors at the Center for Artificial Intelligence in Medicine & Imaging at Stanford University, Palo Alto, California as a post-doctoral research fellow.