Developing a common, international template for documenting and reporting data from major trauma patients

A proposed core dataset and evaluations of its feasibility and reliability

Kjetil Gorseth Ringdal

Dissertation for the degree of philosophiae doctor (PhD)
Faculty of Medicine, University of Oslo
Norway, 2012
© Kjetil Gorseth Ringdal, 2012

Series of dissertations submitted to the Faculty of Medicine, University of Oslo No. 1313

ISBN 978-82-8264-345-0

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub.
The thesis is produced by Unipub merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.
“We can't solve problems by using the same kind of thinking we used when we created them.”

(Albert Einstein, 1879–1955, German-born U.S. theoretical physicist and humanist)
# CONTENTS

1. PREFACE ........................................................................................................ 7
2. LIST OF PAPERS ............................................................................................ 9
3. ABBREVIATIONS ............................................................................................ 10
4. INTRODUCTION ............................................................................................. 11
   4.1 Epidemiology of injuries ........................................................................ 11
   4.2 Benchmarking the quality of care .......................................................... 11
   4.3 The history and benefit of trauma registries .......................................... 12
   4.4 Outcome prediction ................................................................................ 13
      4.4.1 The TRISS methodology ............................................................... 13
      4.4.2 Requirements for the use of prediction models ............................... 14
   4.5 Injury panorama and trauma system mix .............................................. 15
   4.6 Important data variables in trauma care prognostication ..................... 16
      4.6.1 Physiology .................................................................................... 16
      4.6.2 Comorbidity ................................................................................ 16
      4.6.3 Anatomic injury scoring ............................................................... 17
         4.6.3.1 Abbreviated Injury Scale ....................................................... 17
         4.6.3.2 Injury Severity Score (ISS) .................................................. 18
         4.6.3.3 New Injury Severity Score (NISS) ........................................ 19
         4.6.3.4 Scales that use the ICD-9-CM as a basis ............................... 21
         4.6.3.5 Variability in injury severity coding ...................................... 21
      4.6.4 Endpoints in trauma care benchmarking ....................................... 22
   4.7 Trauma registries internationally ............................................................ 23
   4.8 Trauma systems ...................................................................................... 26
   4.9 Variability in trauma registration ........................................................... 27
   4.10 Efforts to standardize data collection in international trauma registries ... 27
   4.11 Efforts to standardize the data collection of European trauma registries ... 27
   4.12 Quality assessment of international datasets ......................................... 28
5. AIMS OF THE PHD PROJECT ....................................................................... 29
6. MATERIAL AND METHODS ......................................................................... 31
   6.1 General introduction ............................................................................. 31
      6.1.1 Feasibility ................................................................................... 31
      6.1.2 Reliability .................................................................................... 31
   6.2 Paper I .................................................................................................... 32
   6.3 Paper II .................................................................................................. 33
   6.4 Paper III ............................................................................................... 34
   6.5 Paper IV ............................................................................................... 35
   6.6 Statistical methods ............................................................................... 36
      6.6.1 Feasibility .................................................................................... 36
      6.6.2 Agreement ................................................................................... 37
      6.6.3 Reliability .................................................................................... 38
   6.7 Ethical approvals ................................................................................... 40
7. RESULTS ....................................................................................................... 41
   7.1 Paper I ................................................................................................ 41
   7.2 Paper II ................................................................................................ 41
7.3 Paper III ........................................................................................................... 42
7.4 Paper IV .......................................................................................................... 43
8. DISCUSSION AND CONCLUSIONS ................................................................... 45
  8.1 Consensus method concerns .......................................................................... 45
  8.2 Core dataset concerns .................................................................................. 46
  8.3 Feasibility concerns ...................................................................................... 48
  8.4 Reliability concerns ...................................................................................... 50
  8.5 Validity concerns .......................................................................................... 52
  8.6 Developing a core dataset and ethical considerations ................................... 52
  8.7 A European trauma registry .......................................................................... 53
  8.8 Challenges met during the projects ............................................................... 53
  8.9 Some conclusions and further perspectives .................................................. 54
9. REFERENCES ...................................................................................................... 57
APPENDIX A ........................................................................................................ 71
APPENDIX B ......................................................................................................... 73
APPENDIX C ......................................................................................................... 74
APPENDIX D ......................................................................................................... 76
1. PREFACE

My interest in emergency medicine and trauma care started when I served in the Medical Battalion of the Royal Norwegian Armed Forces, and when I worked as a registered nurse in the emergency department and in the ambulance services. The interest was further stimulated during medical school in Denmark and during my years as an instructor at the Clinical Skills and Simulation Centre at the University of Southern Denmark, Odense.

Professor Hans Morten Lossius served as my principal advisor and deserves my sincerest gratitude more than anyone else. His visionary entrepreneurship qualities are admirable, and his never-ending optimism, clear mind, and constant support during the project were very important. I am very grateful for being introduced to trauma and emergency medicine research and to his network of scientific clinicians. I am indebted to Professor Petter Andreas Steen, my subsidiary advisor, for his constant support, constructive and timely criticisms, and invaluable contributions in the planning and writing processes. I have been very fortunate to work with such a skilled and highly renowned scientist. He was always available, and his response time was exceptionally short.

Professor Olav Røise, my subsidiary advisor, is acknowledged for providing opportune remarks and very important contributions during the PhD process. His enthusiasm and interest in trauma registry research inspired me considerably. I am also very grateful to him for inviting me to assist in the development of the National Trauma Registry.

I am indebted to Nils Oddvar Skaga, my subsidiary advisor, for sharing his vast knowledge and excellent skills in trauma registries and trauma registry research methodology. His availability, constant support, and constructive and timely criticisms were of great value. I have been very fortunate and privileged to work with him. The structure of his doctoral thesis served as a model for my own.
Senior Lecturer J. Mary Jones is deeply acknowledged for her valuable contributions to study designs and statistical analyses. It has been an honor and pleasure to collaborate with such a competent biostatistician. Her assistance was of great value to the success of papers III-IV.

Senior Researcher Jo Røislien is acknowledged for assisting me with Paper IV. His knowledge of statistical methods and research methodology was very inspiring.

I am also very grateful to nurse anesthetist, trauma coder, and co-author Morten Hestnes for his invaluable help and support in the PhD projects and for comments on the Utstein Trauma Template. His competence, thoroughness, and inquiring nature were inspiring.

I am very grateful to Torsten Eken for his support, for the help I received with Paper III, and for his assistance with the development of the Utstein Trauma Template Dictionary.

I am grateful for the support and the fruitful discussions with my friends, colleagues, and co-authors Thomas Kristiansen, Andreas J. Krüger, Marius Rehn, and Professor Kjetil Søreide. I am proud to have been able to work with such fine scientific clinicians.

Co-authors Professor Timothy J. Coats, Stefano Di Bartolomeo, Professor David J. Dries, Lauri Handolin, Jens M. Lauritsen, Rolf Lefering, and Cameron S. Palmer are acknowledged for their constant support and encouragement.

I am deeply indebted to the Norwegian Air Ambulance Foundation and its nearly 800,000 members who funded my PhD.

Last but not least, the warmest and deepest acknowledgments are extended to my family who provided invaluable support and encouragement during toward the PhD. Without them, I could not have kept the heart warm and head cool throughout the project.

Oslo, March 30, 2012
2. LIST OF PAPERS

This thesis is based on the following five papers; in the text, they will be referred to by their Roman numerals:

Paper I
Feasibility of comparing core data from existing trauma registries in Scandinavia. Reaching for a Scandinavian Major Trauma Outcome Study (MTOS).
K. G. Ringdal, H. M. Lossius; The SCANTEM ad hoc group on Scandinavian MTOS and Trauma Registry.

Paper II
The Utstein template for uniform reporting of data following major trauma: a joint revision by SCANTEM, TARN, DGU-TR and RITG.

Paper III
Collecting core data in severely injured patients using a consensus trauma template: an international multicentre study.
Kjetil G. Ringdal, Hans Morten Lossius, J. Mary Jones, Jens M. Lauritsen, Timothy J. Coats, Cameron S. Palmer, Rolf Lefering, Stefano Di Bartolomeo, David J. Dries, Kjetil Søreide; The Utstein Trauma Data Collaborators.

Paper IV
Abbreviated Injury Scale: not a reliable basis for summation of injury severity in trauma facilities.
Submitted to Injury.
3. ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAAM</td>
<td>Association for the Advancement of Automotive Medicine</td>
</tr>
<tr>
<td>ACS</td>
<td>American College of Surgeons</td>
</tr>
<tr>
<td>AIS</td>
<td>Abbreviated Injury Scale</td>
</tr>
<tr>
<td>ASA-PS</td>
<td>American Society of Anesthesiologists Physical Status Classification</td>
</tr>
<tr>
<td>TR-DGU</td>
<td>TraumaRegister der Deutschen Gesellschaft für Unfallchirurgi (Trauma Registry of the German Society of Trauma Surgery)</td>
</tr>
<tr>
<td>ED</td>
<td>Emergency Department</td>
</tr>
<tr>
<td>ETRN</td>
<td>European Trauma Registry Network</td>
</tr>
<tr>
<td>EuroTARN</td>
<td>European Trauma Audit &amp; Research Network</td>
</tr>
<tr>
<td>GCS</td>
<td>Glasgow Coma Scale</td>
</tr>
<tr>
<td>HAC</td>
<td>Hierarchical Agglomerative Clustering</td>
</tr>
<tr>
<td>ICC</td>
<td>Intraclass Correlation Coefficient</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Disease</td>
</tr>
<tr>
<td>ICD-9-CM</td>
<td>International Classification of Diseases, Ninth Revision, Clinical Modification</td>
</tr>
<tr>
<td>ISS</td>
<td>Injury Severity Score</td>
</tr>
<tr>
<td>ITACCS</td>
<td>International Trauma Anesthesia and Critical Care Society</td>
</tr>
<tr>
<td>LoA</td>
<td>Limits of Agreement</td>
</tr>
<tr>
<td>MTOS</td>
<td>Major Trauma Outcome Study</td>
</tr>
<tr>
<td>NISS</td>
<td>New Injury Severity Score</td>
</tr>
<tr>
<td>NGT</td>
<td>Nominal Group Technique</td>
</tr>
<tr>
<td>NTDB</td>
<td>National Trauma Data Bank</td>
</tr>
<tr>
<td>OUH</td>
<td>Oslo University Hospital</td>
</tr>
<tr>
<td>OUH-U</td>
<td>Oslo University Hospital - Ullevål</td>
</tr>
<tr>
<td>RR</td>
<td>Respiratory Rate</td>
</tr>
<tr>
<td>RITG</td>
<td>Registro Intraospedaliero Multiregionale Traumi Gravi (Italian Trauma Registry)</td>
</tr>
<tr>
<td>RTS</td>
<td>Revised Trauma Score</td>
</tr>
<tr>
<td>SBP</td>
<td>Systolic Blood Pressure</td>
</tr>
<tr>
<td>SCANTEM</td>
<td>Scandinavian Networking Group for Trauma and Emergency Management</td>
</tr>
<tr>
<td>TARN</td>
<td>Trauma Audit &amp; Research Network</td>
</tr>
<tr>
<td>TRISS</td>
<td>Trauma Score Injury Severity Score</td>
</tr>
<tr>
<td>U.S.</td>
<td>United States</td>
</tr>
<tr>
<td>VSTORM</td>
<td>Victorian State Trauma Outcomes Registry and Monitoring Group</td>
</tr>
<tr>
<td>VSTR</td>
<td>Victorian State Trauma Registry</td>
</tr>
<tr>
<td>VSTS</td>
<td>Victorian State Trauma System</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
4. INTRODUCTION

4.1 Epidemiology of injuries

Injury is the leading cause of death for persons aged 5-44 years in high-income countries and the leading cause of death and disability for all age groups <60 years old worldwide. It is estimated that 5.8 million people, of whom two-thirds were men, died from injury in 2004\textsuperscript{1,2}, which equaled 10% of the world’s deaths. This number is understated because ten-fold more people suffer from disability and reduced quality of life due to injury\textsuperscript{2,3}.

In the World Health Organization (WHO) European Region, approximately 800,000 patients died from violence and unintentional injury in 2002, which was approximately 10% of all deaths\textsuperscript{4} and included nearly 42,000 children and adolescents\textsuperscript{5}. Of particular note, the death rates of children in poorer countries were three-times higher than those in richer countries\textsuperscript{5,6}. Additionally, the number of injuries of all ages varies greatly also within Europe, with the lowest number being 26 in the Netherlands and the highest being 151 in Lithuania per 100,000 inhabitants per year\textsuperscript{6}. The differences between and within countries indicate a high potential for reducing injury mortality in certain areas.

While primary injury prevention is probably the most cost-effective approach to this problem, health care systems also have an obligation to monitor and improve the quality of the care that is delivered to trauma victims\textsuperscript{7,8}.

4.2 Benchmarking the quality of care

Benchmarking is an important performance improvement tool that has been used for centuries\textsuperscript{9}. The healthcare industry has recently begun to use this tool more systematically, thus increasing the potential to significantly improve the efficiency, cost-effectiveness, and quality of care. Benchmarking is a method for comparing processes and performance metrics to those of other hospitals or systems and
identifying “best practices”. Dimensions that are typically measured are quality, time, and cost. The results and processes of a service can be evaluated year-by-year (internal benchmarking) or can be compared to other services (external benchmarking) with a defined set of quality indicators. The greatest value to be gained from the performance data that health care facilities collect may emerge from processes that compares such data. Regional or national benchmarks can identify and reduce variation and increase quality, and collaborative benchmarking can help organizations go beyond data comparison to understand the key processes that contribute to the effective delivery of health care.

Data are typically collected in registries, and valid and reliable comparisons require comparable data with common terminology, data definitions, and quality indicators.

4.3 The history and benefit of trauma registries

The modern era of trauma registries appears to have begun in 1969 with a computerized registry at Cook County Hospital in Chicago, Illinois. This registry served as a prototype for the multi-hospital Illinois Trauma Registry that began operation in 1971. Trauma registries were originally designed to document the acute phase of hospital treatment that was delivered to the most seriously injured patients. These registries can also be used for outcome-based benchmarking using outcome prediction models and for process-based benchmarking (e.g., care process efficiency) or resource-efficiency of treatment. In addition to the primary aim of enabling comparative analyses to improve the quality and optimal care of injured patients, trauma registries constitute a major source for injury prevention research and can be used to generate hypotheses, plan protocols, or identify potential study participants for clinical trials. Trauma registries can clarify whether the rapid transport of trauma victims is preferable to more advanced interventions on site, to evaluate pre-hospital field triage protocols, and to measure
the potential advantage of helicopters over ground transport\textsuperscript{7, 25}. Trauma registries can also be used to evaluate the post-hospital phase of treatment, such as rehabilitation institutions and ambulatory physiotherapy, but most include little information beyond the discharge destination\textsuperscript{7}. Unfortunately, most trauma registries, especially those that are hospital-based, are not designed to track the patient course and outcome through the entire trauma-care chain but are limited to one part of the chain, such as the acute care hospital\textsuperscript{24, 26, 27}.

4.4 Outcome prediction
Prediction models are commonly used to calculate the probability of survival of a severely injured patient and to identify “unexpected deaths”\textsuperscript{28}. Crude mortality rates that were previously used as indicators of the quality of care are of limited value if they are not adjusted for the patient risk profile\textsuperscript{29}. Risk stratification models attempt to increase the reliability of predicting the relationship between patient and injury characteristics to observed outcomes\textsuperscript{30}. The aim of risk adjustment is to identify characteristics that are strongly related to mortality and that could, without proper attention, confound analysis. This adjustment is made to correct for case-mix variations that may affect the outcome independent of the quality of care. Without risk stratification, trauma centers that treat high-risk patients (e.g., with severe anatomical injury) will appear to have worse results than others\textsuperscript{31}. Such medical prognostic models are used in various settings and for various reasons, such as to inform about the future course of the patients after injury or to evaluate the observed outcome against the expected outcome in a benchmarking procedure against a recognized standard or versus other hospitals\textsuperscript{32, 33}.

4.4.1 The TRISS methodology
Over the last two decades, the United States (U.S.) Trauma Score Injury Severity Score (TRISS) outcome prediction model\textsuperscript{26, 34} has been the most commonly used
prognostic model for comparing outcomes of trauma patients. The TRISS model adjusts for differences in the severity of the anatomic injuries, physiological derangements, the mechanism of injury, and age. Physiological derangement is based on the patient’s Revised Trauma Score (RTS) components, while anatomic injury severity is scored using the Injury Severity Score (ISS). The original TRISS coefficients for outcome prediction were derived from the U.S. Major Trauma Outcome Study (MTOS) but have recently been updated based on patient cases from the National Trauma Data Bank (NTDB). While TRISS has several limitations and has been criticized by many authors, it continues to be the most widely accepted tool for comparisons of trauma outcomes worldwide.

4.4.2 Requirements for the use of prediction models
A prognostic model should have clinical credibility, be accurate (well-calibrated and with good discriminative ability), be generalizable (externally validated), and, ideally, be shown to be clinically effective, i.e., provide additional clinical information that enables improved treatment decision-making and patient outcome. The use of prognostic models requires unambiguous definitions of predictors and outcome variables. As described above, a prognostic model may not be extrapolated to other systems with a different case-mix unless it includes all of the important prognostic variables and those variables are appropriately modeled. The difficulty is to know whether a model actually includes all important variables. However, one way of reducing case-mix variation would be to define a set of patient inclusion and exclusion criteria for a registry or a prognostic study. Furthermore, an evaluation of the degree of dissimilarity between the patients in different centers (known as variation in case-mix) and of selection criteria must be undertaken before transporting a prognostic model to a different trauma system or trauma population.
4.5 Injury panorama and trauma system mix

European registry publications indicate that the proportion of penetrating trauma (i.e., injuries resulting from tissue penetration or puncture by a sharp object\textsuperscript{55}) varies between 1.5% to 12%, with the largest 11-nation European Trauma Audit & Research Network (EuroTARN) study from 2007 reporting 4%\textsuperscript{48, 56-59}. The U.S. MTOS study from 1990 reported a higher fraction of traumas (21%)\textsuperscript{38}, while more recent U.S. data, including the 2010 report from the NTDB, indicated a lower, 8-10% incidence\textsuperscript{60, 61}. The great variability in reported fractions may be real but may, at least partly, be due to varying definitions of trauma, various mechanisms of injury, inclusion criteria, or data capture methods. To properly identify the proportion of penetrating trauma, common definitions should be agreed upon.

It is known that differences exist between the European and North American trauma care system organizations, and that these differences could lead to case-mix variation and selection differences. In many U.S. systems, paramedics and emergency medical technicians work to minimize pre-hospital time (the ‘load and go’ principle)\textsuperscript{62}, while in other systems, particularly in some European countries, specialized pre-hospital physicians initiate first emergency interventions, such as endotracheal intubations, before transporting the patient to the hospital (the ‘stay and play’ principle)\textsuperscript{37, 63}. Therefore, hospital admission data may be difficult to compare because the U.S. TRISS model excludes patients that have been intubated under general anesthesia on the scene. In contrast, these patients form a significant proportion of the European trauma population\textsuperscript{64}.

Furthermore, emergency department (ED) deaths are handled differently in the U.S. and many European countries. In Norway and in the EU, a patient who die in the ED is per definition admitted\textsuperscript{65}. North American patients are often admitted to the hospital only upon leaving the ED, i.e., a patient dying in the ED is not hospitalized\textsuperscript{66-69}. Hence, uncertainty exists as to how U.S. trauma centers report deaths prior to hospital admission. This administrative variability provides difficulties for trauma registries that must be specifically addressed\textsuperscript{37}.
4.6 Important data variables in trauma care prognostication

4.6.1 Physiology
In the TRISS model, the physiologies of the central nervous, circulatory, and respiratory systems after injury are reflected by the RTS variables, i.e., the Glasgow Coma Scale (GCS) score\textsuperscript{70}, systolic blood pressure (SBP), and respiratory rate (RR). The current TRISS convention requires RTS variable scoring upon ED admission. In many European countries, especially those that favor the Franco-German model of emergency medical service care\textsuperscript{71, 72}, which aims to “bring the hospital to the patient”, many head injury and other severely injured patients are anesthetized and intubated before ED arrival by specially trained, pre-hospital, critical care physicians\textsuperscript{73-76}. As GCS scores and RR values from the ED are not attainable from these patients, large amounts of data may be missing for these patients\textsuperscript{48}. In Europe, these patients are only assigned GCS and RR scores recorded immediately before intubation, either on scene of the injury, or in the ED of a local referring hospital.

4.6.2 Comorbidity
Pre-existing medical conditions and physical status (comorbidity) can influence the outcomes of severely injured patients\textsuperscript{77-81}, and Morris et al. suggest an association between the level of physiological reserves prior to trauma and the outcome of the patient\textsuperscript{77}. Some authors argue that pre-existing medical conditions should be included in trauma survival prediction models\textsuperscript{81-83}, while others have failed to find that this information increases the precision of the models\textsuperscript{84, 85}. In various studies, pre-existing diseases have been dichotomously classified as present or absent\textsuperscript{82, 86-88}, as an ordinal variable that is calculated from the number of pre-existing International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-
9-CM) disease codes\textsuperscript{77, 78, 89-93}, or according to the Charlson Comorbidity Index\textsuperscript{94-97}, which also takes into account the number and seriousness of such diseases. An alternative to the above methods of classification is categorizing pre-injury comorbidity according to the American Society of Anesthesiologists Physical Status (ASA-PS) system\textsuperscript{81, 98}, which estimates the global physiological impact of pre-existing diseases on a patient, regardless of whether this impact is caused by one or several disease processes. Skaga et al.\textsuperscript{81} reported that pre-injury ASA-PS\textsuperscript{98} was an independent predictor of survival in trauma patients when also adjusting for the traditional TRISS model variables\textsuperscript{81}. Pre-trauma comorbidities are not accounted for in any of the present outcome prediction models, a fact that should be considered when benchmarking with case-mixed variation.

4.6.3 Anatomic injury scoring

Injury classification by type and severity is considered fundamental to trauma research and quality control assessments\textsuperscript{7-99}. Worldwide, classification systems of the severity of trauma patient’s anatomic injuries are based either on Abbreviated Injury Scale (AIS) codes\textsuperscript{100} or on scales based on ICD-9-CM codes\textsuperscript{101}.

4.6.3.1 Abbreviated Injury Scale

The AIS system, which was initially developed by the U.S. Committee on Medical Aspects of Automotive Safety in 1971\textsuperscript{102} and further developed by the U.S. Association for the Advancement of Automotive Medicine (AAAM)\textsuperscript{103}, is most frequently used to classify the severity of anatomic injuries in North America, Europe, Australia, New Zealand, and Japan\textsuperscript{104}. It is an anatomically based, consensus-derived scoring system that enables the comparison of morphologically different injuries of similar severity in studies involving multiple patients\textsuperscript{100}. The first six AIS-code digits designate the injured body region, the general type of anatomic structure, the specific anatomic structure, and the nature of the injury.
The seventh digit refers to an ordinal injury severity scale that ranges from 1 (minor injury) to 6 (maximal injury)\textsuperscript{100}. The AIS dictionary has nine chapters, which correspond to nine body regions. The AIS codes form the basis for several summative injury severity scores, such as the ISS and the New Injury Severity Score (NISS)\textsuperscript{105}.

The AIS system has several limitations and has been widely criticized\textsuperscript{7, 106-111}. Limitations to this system include the following: the scale does not have biunique intervals (i.e., the relationship between AIS scores and mortality is not linear), the relationships between the survival rate and maximum AIS by the body region of the patients vary, and the difference between AIS 1 and 2 is not the same as that between AIS 4 and 5\textsuperscript{106, 112}. For example, Copes et al. found that the survival rate for patients whose most serious injury was an AIS 3 to the head/neck was considerably lower than for patients whose most serious injuries were an AIS 3 in other body regions\textsuperscript{106}.

Despite its limitations, the AIS system continues to be the most accepted and widely used quantitative tool for classifying and describing injuries\textsuperscript{104, 113}.

### 4.6.3.2 Injury Severity Score (ISS)

The ISS was developed in the U.S. in 1974\textsuperscript{97} and remains the most popular score for quantifying the impact of anatomic injury. It classifies and quantifies the overall severity of injury across body regions and is calculated by summing the squares of the highest AIS severity codes in each of the three most severely injured of the six ISS body regions (see patient example in Appendix C). ISS scores range from 1 to 75, and by convention, AIS 6 injuries give an ISS score of 75, regardless of other injuries. The ISS includes the following weaknesses: any error in AIS scoring increases the ISS error, many different injury patterns can yield the same ISS score, and each body region is weighted equally (i.e., the additional importance of head injuries in trauma mortality is ignored). Because ISS only accounts for the most
severe injuries in the three most severely injured ISS body regions, additional injuries will not be represented\textsuperscript{106, 114}. Furthermore, although the ISS is scaled from 1 to 75, it actually only assumes 44 discrete values. As most ISS values fall within the lower end of the scale, these values are not uniformly distributed\textsuperscript{112, 114, 115} nor is ISS a monotone scale\textsuperscript{112}. Additionally, the mortality rate of a patient with an ISS of 25 could be higher than the mortality rate of a patient with an ISS of 27 due to the different AIS score combinations that comprise the ISS score\textsuperscript{112} (Figure 1). Although some authors claim that ISS is ordinal\textsuperscript{104, 112, 116}, it is probably more correct to treat ISS as a “nominal-like” scale.

![ISS vs. Mortality](image)

**Figure 1.** ISS vs. Mortality for 10,724 patients from the OUH-U Trauma Registry. The endpoint is 30-day mortality.

**4.6.3.3 New Injury Severity Score (NISS)**

NISS was introduced in 1997\textsuperscript{105}. This score is a revised version of the ISS but is more
easily calculated as the sum of squares of the three most severe AIS injuries regardless of body region (see patient example in Appendix C). Therefore, this scoring better prioritizes the most severe injuries.

NISS has shown better discrimination (i.e., the ability to distinguish between patients who experience the outcome of interest and those who do not\textsuperscript{50, 117}) and calibration (i.e., the ability to predict an outcome that closely reflects the true outcome\textsuperscript{50, 117}) than ISS\textsuperscript{105}. A NISS score will be equal to or greater than an ISS score for any given patient, and this type of scoring appears to be more accurate for summarizing injury severity\textsuperscript{118, 119}, which is particularly evident for patients with multiple traumatic brain injuries or a penetrating injury to one ISS body region\textsuperscript{120}. Although NISS avoids many of the acknowledged limitations of ISS, it also only takes into account a maximum of three injuries, and like the ISS, it is not linearly related to mortality\textsuperscript{112, 121} (Figure 2). As described above for ISS, the mortality rate for a patient with a NISS of 25 is higher than that for a patient with a score of 27\textsuperscript{112}. 

![NISS vs. Mortality](image-url)
4.6.3.4 Scales that use the ICD-9-CM as a basis

The ICD-9 Injury Severity Score (ICISS) is based on the ICD-9-CM codes. It calculates the survival risk ratio (SRR) for each ICD-9-CM code and uses the product of each SRR to form an overall injury severity score, ICISS99,122. A given ICD-9 SRR represents the likelihood that any individual patient will survive the particular ICD-9 injury. Several studies have shown that SRR values that were derived from ICD-9 codes are as good as or better than ISS scores as predictors of survival122-124. However, because ICISS is often based on ICD codes from discharge notes in hospital administrative data and because ICISS is implicitly dependent on all factors that contribute to patient mortality, including the care process99, its accuracy has been questioned. Additionally, because this system has only been developed for ICD-9, it is not applied in many European countries where the ICD-10 edition is used.

In 1985, MacKenzie et al. developed the ICDMAP-85125, which converts ICD-9-CM codes to AIS codes, and an updated version, the ICDMAP-90 was released in 1997126, 127. In comparative studies, ICD-9-to-AIS mapped scores generally did not perform as well than those based on directly coded AIS scores or the ICISS system126-131.

4.6.3.5 Variability in injury severity coding

No uniform approach to the classification of anatomic injury severity has been agreed upon in Europe or worldwide, and individual trauma registries use various versions of AIS or ICD-9-CM-based systems16.

The NTDB data were, until 2007-2009, not standardized; therefore because the data that had been collected from different trauma-receiving facilities differed in
meaning, the data were not comparable\textsuperscript{132}. Recently, NTDB has developed a data dictionary (i.e., the National Trauma Data Standard)\textsuperscript{133} for a more uniform data collection. However, trauma-receiving hospitals are currently not required to send AIS codes to the NTDB, and there are no AIS version requirements. Furthermore, hospitals use a variety of AIS versions (i.e., AIS 1980 to 1998 editions) and do not code AIS consistently\textsuperscript{60, 134}. For these reasons, ISS values that are used in report analyses are calculated with AIS codes that have been submitted by hospitals and then cross-walked to the AIS 1998 edition codes. If a hospital does not submit AIS codes, the ISS is based on AIS codes that are generated from ICD9-CM codes using the ICDMAP-90 mapping program\textsuperscript{60}. While the system is imperfect, comparisons are considered valid because it is consistent across centers [personal communication with the NTDB].

4.6.4Endpoints in trauma care benchmarking

Survival status is a principal outcome measure in many biomedical fields. In trauma care, hard, short-term endpoints are typically used to benchmark outcomes and the quality of care. The U.S. TRISS model uses the variable ‘survival at end of acute care stay’, which has several quality assessment limitations. Several outcome variables are currently applied in Europe and worldwide: ‘end of acute care stay’, ‘30-day outcome’, ‘30-day in-hospital outcome’, and ‘90-day mortality’ to name but a few\textsuperscript{16, 24, 64, 135, 136}. Different outcome definitions will introduce mortality indicator biases in trauma care comparisons. Outcomes measured by in-hospital survival can be particularly problematic in cases where patients remain in the hospital for shorter lengths of time and where patients are transferred between facilities at a greater frequency. Thus, a greater proportion of deaths that occur within 30 days of injury may be missed if only in-hospital deaths are considered\textsuperscript{24, 135, 137-139}. Death occurring later than 30 days after injury is more likely caused by other conditions, such as a pre-existing disease\textsuperscript{135, 140}. The Trauma Audit & Research
Network (TARN) showed that 4.8% of 69,650 patients died within 93 days of admission\textsuperscript{64}. Of these patients, only 9% died later than 30 days after admission; and these deaths mainly included those with a low ISS (<9) that had an age >65 years. In a Norwegian trauma registry, 4.6% of deaths occurred later than 30 days after injury\textsuperscript{135}, whereas in-hospital 30-day mortality data from the German trauma registry (cases with NISS ≥16) indicate that 4.9% of the patients who died did so later than 30 days after being injured [personal communication with the DGU-TR]\textsuperscript{24}.

The lack of international consensus on the definition of outcome after trauma is remarkable and represents a serious methodological issue. The U.S. survival endpoint ‘end of acute care’ is not clearly defined, has only recently been criticized in the U.S.\textsuperscript{67, 97, 136, 141-144}, and may invalidate U.S. prediction models in European trauma care.

4.7 Trauma registries internationally

Several trauma registries exist throughout the world, and some of the most important registries are described below.

United States of America

The National Trauma Databank (NTDB), which is run by the American College of Surgeons (ACS) Committee on Trauma, is an aggregation of U.S. and Canadian trauma registry data and is the largest existing trauma registry data bank\textsuperscript{145}. This database is a continuation of the MTOS data collection initiative from 1982 through 1987\textsuperscript{38}. NTDB aims to improve the care of injured patients through systematic efforts in prevention, care, and rehabilitation, and to inform the trauma community, the public, and decision makers about a wide variety of issues that characterize the state of care for injured persons. In the U.S., the ACS requires that all certified, Level I, II, III and IV trauma centers\textsuperscript{146, 147} have a trauma registry\textsuperscript{23}. To
date, the databank contains more than five million patient records. Nearly half of these patients have suffered minor injuries, and under one-fourth have suffered severe injuries.

**United Kingdom (UK)**

The TARN\textsuperscript{148} was established in 1990 to perform clinical audits and research in trauma care. It provides evidence nationally and has become the national auditor of trauma care, presently collecting data from 180 trauma-receiving National Health Service (NHS) Trusts hospitals across England and Wales. To date, the database contains approximately 320,000 cases [personal communication with TARN]. TARN monitors the standards of trauma care that were set out by the Royal College of Surgeons, British Orthopaedic Association, and National Institute for Health and Clinical Excellence (NICE) and provides each NHS Trust and Commissioner with case-mix-adjusted outcome analyses and comparisons of trauma care across institutions. The main objective of this registry is to facilitate the development and improvement of trauma services, thereby reducing the associated burden of death and disability\textsuperscript{149, 150}. The organization is funded through membership fees that are paid by participating Trusts.

**Germany**

The Trauma Registry of the German Society of Trauma Surgery (TraumaRegister DGU\textsuperscript{\textregistered} der Deutschen Gesellschaft für Unfallchirurgie) (TR-DGU)\textsuperscript{151} was originally set up in 1993 to enable multicenter collections of data from severely injured patients in the German speaking area\textsuperscript{152}. Participation has increased steadily, and currently, 370 hospitals from Germany, Austria, Switzerland, Netherlands, and the United Arab Emirates submit data to this registry [personal communication with TR-DGU]. Hospitals use slightly different selection criteria, but quality/mortality analyses have, until recently, only included patients that receive initial treatment and who have an ISS above 15\textsuperscript{21, 153}. To date, the registry contains more than 50,000
severely injured patient incidents\textsuperscript{56}.

\textit{Australia}

In Australia, the first truly state-wide trauma registry was established in Victoria in 2001\textsuperscript{154}. The Victorian State Trauma Registry (VSTR)\textsuperscript{155}, which is managed by the Victorian State Trauma Outcomes Registry and Monitoring Group (VSTORM)\textsuperscript{156}, monitors the performance and effectiveness of the Victorian State Trauma System (VSTS) and collects information on all major trauma patients from all health care facilities that manage major trauma patients in Victoria. The VSTS aims to reduce preventable death and permanent disability and to improve patient outcomes by having the right patient delivered to the right hospital in the shortest amount of time\textsuperscript{157}. The VSTS has overseen the following significant improvements in the care of major trauma patients: 80\% of all major trauma patients are treated at a major trauma facility, mortality rates have been reduced with fewer lower than the expected deaths according to international benchmarks, positive trends in preventable deaths have occurred, and patients now have reduced lengths of stay in hospitals\textsuperscript{157}.

Efforts have recently been initiated to establish national trauma registries in Australia and New Zealand\textsuperscript{158}.

\textit{Injury epidemiology databases in Europe}

In Europe, epidemiologic injury data are publicly accessible through international data providers such as Eurostat\textsuperscript{159}, the EU Injury Database\textsuperscript{6}, and the WHO Europe\textsuperscript{160}. These databases are not used for routine monitoring of the quality of care using risk stratification models and medical quality assessment approaches to the same extent as the other databases. Rather, they use more epidemiological and public health approaches\textsuperscript{161}. 
4.8 Trauma systems

Formalized systems of trauma care, which were first described more than three decades ago in the U.S.\textsuperscript{162}, constitute structured, core operational components of care processes from the site of injury, through hospital care, to rehabilitation\textsuperscript{37}. Supported by an increasing amount of empirical evidence, the benefit of trauma systems has been widely accepted among health care providers\textsuperscript{21, 163-166}.

Trauma systems can be either inclusive or exclusive in nature. Inclusive trauma systems incorporate high-level trauma centers for the most severely injured patients and acute care facilities for less severely injured patients. An exclusive trauma system is organized around a high level (I) trauma center and mainly cares for severely injured patients\textsuperscript{164, 167, 168}. In more remote and rural areas, inclusive trauma systems in which every designated hospital is mandated to assess and stabilize patients before transferring them can improve care. In densely populated urban areas, exclusive trauma systems provide readily available, trained trauma personnel to deliver definitive care without the potential delays of an inclusive system\textsuperscript{164, 168}.

In Europe, trauma systems are the exception rather than the rule\textsuperscript{169}. Germany, the country that has reached the furthest with developing trauma systems\textsuperscript{153, 170-174}, has documented a reduction in mortality after a regionalization of trauma care (trauma networks)\textsuperscript{21, 63}. In the UK, the recent development of London’s trauma system gives reason to expect promising results in the future\textsuperscript{175, 176}.

An improved understanding of the benefits and limitations of different trauma care settings and systems requires the comparison of these organizations across regional and national systems\textsuperscript{21, 177}. Such studies have been few and largely limited to North America\textsuperscript{95, 166}, but Australian and European researchers have recently increased the focus on this important aspect of quality improvement\textsuperscript{19, 177-181}.
4.9 Variability in trauma registration

An adequate evaluation of the global burden of injury and a meaningful evaluation of outcomes require appropriate data collection tools, such as trauma registries, which should utilize uniform variables to enable valid comparisons. As described above, some countries document trauma systematically but use varying datasets and few compatible definitions of common variables. This lack of uniformity poses substantial challenges to most initiatives seeking to assess the quality of healthcare systems across regions. Using the same language and a common set of data variables and definitions is the first step in benchmarking trauma care systems.

4.10 Efforts to standardize data collection in international trauma registries

In 1999, a working group from the International Trauma Anesthesia and Critical Care Society (ITACCS) published an Utstein-style template with recommendations for the uniform reporting of data following major trauma. This template focused on data from the pre-hospital phase, early in-hospital management, co-morbidity, and outcome. In accordance with previous Utstein templates (particularly for cardiac arrest), variables were classified as “core” (essential) or “optional” (supplemental). However, despite the intention of facilitating uniform data collection and reporting, few publications are based on this template, indicating that further development, with a major reduction in the large number (92) of core variables and more precise definitions is needed.

4.11 Efforts to standardize the data collection of European trauma registries

The difference between European health care systems and the pace of integration between the health care systems is slow because of different social, economic, geographical, cultural, and political barriers. These obstacles result in differences in the structure, organization, and focus of healthcare systems for trauma care. There
has also been some reluctance in sharing local and national data, although it has been recognized that lessons learned in one area of Europe may be useful for others\textsuperscript{59}.

To re-address the need for a core dataset, members of the European trauma registries met in 2002. This collaboration, the EuroTARN Group, assessed the potential for creating a data collection trial among a number of trauma services and for the exchange and comparison of summative data and crude mortality rates. Although severity-adjusted outcome comparisons were not possible at that time because of registry differences\textsuperscript{59}, it was possible to share summative data on a European level, and there was great interest from several countries in sharing their trauma care experiences.

It is assumed that a European trauma database can reveal trends in care, find patient, hospital, or system outliers, and assist hospitals and systems that require more attention to increase the quality of care at these hospitals. In other words, this type of database use would develop systems that could use European diversity to identify key factors that are related to good outcomes, thereby building better trauma systems by utilizing the best parts of many different systems\textsuperscript{16, 24, 161}. Implementing these across systems should reduce the process and levels of service variations and allow all participants to learn and improve their organizations\textsuperscript{9}.

4.12 Quality assessment of international datasets

Although many countries and regions have established trauma registries, there have been few published, true dataset evaluations on the levels of feasibility, reliability, and validity. To my knowledge this is a common problem in trauma registry research.
5. AIMS OF THE PHD PROJECT

The purpose of this thesis was to develop a common international template for documenting and reporting the care of severely injured patients, assess its feasibility, and test the reliability of some of its important variables. The aims of the four individual papers were the following:

I. To evaluate whether using data variables from existing major Scandinavian trauma registries for a joint, prospective Scandinavian MTOS was feasible and to describe the differences and similarities between registries, emphasizing structure, inclusion and exclusion criteria, data variable definitions, and common data variables.

II. To carry out a consensus process between major European trauma registries to further develop the Utstein Template for Uniform Reporting of Data following Major Trauma by making it compatible with the trauma registries and adhering to EuroTARN. Accomplishing this aim should support the establishment of a European Trauma Registry and promote further development of a European outcome prediction model for auditing and benchmarking.

III. To test the feasibility of collecting patient-level data for severely injured patients from trauma centers using the revised Utstein Trauma Template variables.

IV. To estimate the levels of agreement and reliability of ISS and NISS, based on AIS 2008, of a representative group of AIS-certified, Norwegian trauma registry coders and compare them to a reference standard.
6. MATERIAL AND METHODS

6.1 General introduction
The aims of the PhD project were reached via a triangulation of methods with both qualitative and quantitative approaches. To evaluate whether data variables from Scandinavian trauma registries were sufficiently similar, database variables and definitions from each registry were collected. To establish a consensus-based, common core dataset, an expert panel was established, and patient-level data from Scandinavian, European, Australian, and U.S. study populations were used to evaluate the feasibility of collecting Utstein Template data. To estimate data variable reliability, we cooperated with trauma registry coders.

6.1.1 Feasibility
Before recommending its use, it is essential to identify the problems in applying a template to clinical data. The feasibility (i.e., ease of use, collectability, and acceptability) should be evaluated for clinicians, data collectors, and the target population\textsuperscript{192}. A pilot trial involving hospitals that are used to collect lots of data, will not usually give a true indication of how easily these variables are collected. A mixture of small, medium, and large hospitals or services with different levels of data collection experience should be included\textsuperscript{192}. In cases of data collection difficulties, data variables, response categories and definitions may be improved or excluded.

6.1.2 Reliability
Uniform dataset variables should be evaluated with regard to reliability (i.e., the extent to which a measurement gives consistent results) by measuring the degree of agreement between raters\textsuperscript{193, 194}. Human judgment can vary, and the same raters may rate data differently. Good agreement and reliability of data when multiple
raters record and score that data from the same subjects implies that the variable is rater-independent\textsuperscript{192}. Unreliable rater scoring will dilute the association of predictors with outcome and is termed “regression dilution bias”\textsuperscript{195}. Reliability places an upper limit on validity so that the higher the reliability, the higher the maximum possible validity\textsuperscript{194}. Because the reliability of a data variable or rating tool influences its validity, a reliability study should be performed prior to a validity study\textsuperscript{192}.

6.2 Paper I

Inclusion and exclusion criteria, core data variables, and data definitions were collected from hospitals and registries that were involved in the establishment of the Scandinavian Networking Group for Trauma and Emergency Management (SCANTEM)\textsuperscript{196}. The aims of SCANTEM at that time were to create a Scandinavian major trauma registry and enable a Scandinavian major trauma outcome study. Data variable comparisons were based on datasheets with core data variables that were received from the participating centers and expanded via unstructured interviews and e-mail correspondence. Based on the information that was received, we identified inclusion and exclusion criteria and common core data and attempted to reveal the degree of uniformity of variable definitions. Examples of how to construct and define data variables were obtained from the data dictionaries of the U.S. NTDB\textsuperscript{133}, TARN\textsuperscript{197}, Norwegian National Trauma Registry (NNTR), the Swedish trauma registry standard (KVITTRA), and the Utstein Style of uniform reporting of data following major trauma\textsuperscript{187}. Core data were defined according to the Injury Surveillance Guidelines (ISG) of the WHO\textsuperscript{11} and the Utstein publication\textsuperscript{187}. Core data from the participating registries were categorized as in the suggested EuroTARN core dataset\textsuperscript{59} because this dataset represented the first step towards a common, European core dataset.
6.3 Paper II

A European expert panel was invited to a nominal group technique (NGT) process\textsuperscript{108, 109}, which was modified to fit our purpose. Potential panelists were identified from the participants of Paper I\textsuperscript{104}, from the previous EuroTARN Group\textsuperscript{59}, through unsystematic searches in PubMed and Google Scholar, and through proposals from invited panel members. The focus of this project was to establish a panel comprised of individuals who were central to developing and managing European trauma registries and organizations. Therefore, clinical experts, database managers, data collectors, biostatisticians, and epidemiology experts, who were all experienced in the fields of trauma registry research and trauma audits, were invited to partake in the project.

The modified NGT process included four steps. First, each panelist received the necessary background documents and was asked to return proposals for inclusion and exclusion criteria and a prioritized list of a maximum of 30 core variables. A limit of 30 variables was utilized to force panelists to focus on the most important data for comparison of care and prediction modeling. It was also assumed that collecting too many variables would affect data quality and make template implementation difficult. The results were summarized and structured by the coordinators and redistributed by e-mail to the panelists.

In the second step, the experts were asked to comment on the suggested data variables and to reconsider and re-prioritize their suggested 30 variables. The third step consisted of two consensus meetings, each with a structured variable re-evaluation, discussion, and conclusion process. In the fourth step, panelists were allowed to give final comments and conclusion approvals by e-mail. Finally, a letter of consent (i.e., a letter of agreement) was signed by all participants to enhance the implementation of the achieved variables in daily trauma registry practice.
6.4 Paper III

In this prospective, international multicenter feasibility study, a mixture of small-, medium-, and large-volume European trauma centers was asked to collect and code up to 50 consecutively hospitalized trauma patients. Some North American and Australian centers were also invited to participate to enable a wider comparison. Registries and centers were identified from Paper I, Paper II, through personal contacts with other known registries, and through unsystematic searches in PubMed and Google Scholar for active registries. Trauma centers were invited using a standardized open letter that was sent by email. Three reminder emails were sent to the centers that had agreed to participate if data had not been submitted by the deadline. Institutions that did not respond to the first invitation were not followed up.

The centers were expected to include consecutive data on trauma patients with a NISS ≥16 between September 1st, 2009 and November 30th, 2009. The exclusion criteria included asphyxia, drownings, burns as predominant injury, interhospital transfers, hospital admission >24 hours after injury, and patients declared dead before hospital arrival or without signs of life upon hospital arrival and no response to initial hospital resuscitation 24,55.

A sample size of 50 cases per center was chosen to balance study workload with a reasonable number for computing completeness proportions. Low-volume trauma centers that were not able to collect 50 cases within the timeframe were asked to submit three-months worth of data.

Participants were asked to collect all Utstein Trauma Template variables55, to fill out a questionnaire on the variables that they were able to collect, to determine whether their local definitions deviated from the template definitions, and to cite what data collection difficulties they experienced. Moreover, additional comments could be made for each variable. The centers were asked to use the whole, seven digit AIS code according to the AIS 2005 or 2005-update 2008100.
The main outcome variables included the following: data completeness at patient and variable levels, variable definition discrepancies, and collection difficulties. For patient-level data, unknown and blank values were considered to be missing values.

Patient-level data were collected from the local trauma registries, and participants without a suitable registry were provided with an electronic database that had been developed by the investigators. Centers that did not return the questionnaires, patient-level data, or a signed consent form were excluded from the study.

6.5 Paper IV

Potential study participants included nineteen Norwegian AAAM-AIS-certified trauma registry coders who were working in trauma registries or who were intended to code in hospitals that were in the process of establishing a registry. The list was cross-checked against a list of Norwegian Better & Systematic Trauma Care Foundation network contact persons.

Patient records were selected from the trauma registry at Oslo University Hospital - Ullevål (OUH-U), which receives approximately 1400 trauma patients annually. According to the database, 34% of the cases had an ISS >15 (i.e., severe injury) and 44% had an NISS >15, according to the AIS 2008 edition. A sample of consecutive patients with a NISS >15, who were directly admitted from the scene of injury, were selected for the study. Exclusion criteria included asphyxia, drownings, and burns as the predominant injuries, hospital admission >24 hours after the injury, and patients declared dead before reaching the hospital and with no signs of life or response to initial resuscitation upon arrival to the ED.

As we wanted conditions resembling the daily work of a clinical trauma registrar, all participants were asked to independently score the cases according to an expanded version of the Utstein Trauma Template, which was comprised of 48
data variables, including the AIS, ISS, and NISS. Finally, the raters were asked to complete a questionnaire on their level of experience.

The main outcome measures included the completeness of injury coding, agreement in ISS and NISS scoring, and the reliability of ISS and NISS scoring. The determination of the sample size for an ISS/NISS reliability study depends upon a reasonable estimate of the reliability coefficient in the study population, the coefficient of the confidence interval, and the maximum error (e.g., from a previous estimate)\textsuperscript{192, 194}. A sample of 50 medical records (cases) has been suggested\textsuperscript{194} and thus used in Paper IV. This consideration was also judged achievable within a reasonable time and at reasonable costs.

Web-based data-entry tools were used to collect the clinical and questionnaire data. A reference panel of three experienced AIS coders developed a reference standard that was based on the AIS 2008 dictionary. Inter-panel disagreements were resolved by consensus or by consultation with clinical experts in the respective surgical fields.

6.6 Statistical methods

Data were analyzed using SPSS version 18 and 19 (IBM Company, Chicago, IL, USA), Stata/SE version 11.2 (StataCorp LP, College Station, TX, USA), and R version 2.11.1 (The R Foundation for Statistical Computing, Vienna, Austria)\textsuperscript{201}.

6.6.1 Feasibility

In Paper III, feasibility was measured for each Utstein data variable by determining the following factors: 1) the center’s self-reported ability to collect the variable, 2) deviances from template variable definitions, 3) variable collection difficulties, and 4) the completeness of Utstein core data from each participating center presented as counts and proportions and with 95% confidence intervals (CI). Completeness was
further judged by the number of centers reaching completeness of each variable at percentile levels of 50, 75, and 100.

The desired goal for data reporting completeness was set at ≥80%, which was based on statistical consultations with Professor Douglas G. Altman of the University of Oxford, UK. Because there are no published discussions on the acceptability levels of missing data in trauma registry studies (e.g., prognostic studies), controlled trials, or follow-up rates, this is essentially an “arbitrary” clinical limit. Completeness was not categorized into groups because such categories would not include confidence limits; and thus, the categories would not show the variation of completeness within each data variable or within participating centers.

6.6.2 Agreement

_Bland and Altman’s limits of agreement_

In Paper IV, inter-rater agreement was assessed using the limits of agreement (LoA) method that was introduced by Bland and Altman^{202,203}. This method compares the estimated variation in the data to a clinical evaluation of what is an acceptable variation in order for measurements to be considered “not different”. The LoA were calculated as the mean of the differences between the measurements of two raters ± 1.96 × standard deviation and contain 95% of future measurement pairs on similar individuals, assuming a normal distribution of data and independent observations^{204}. A smaller range between these two limits denotes better agreement^{205} (Figure 1).
Figure 1: Bland-Altman plot of the differences against the mean.

An extension of the LoA method to compare more than two pairs of measurements (i.e., raters) was used in this analysis. A clinically acceptable LoA range was set at ±9 units, which was equivalent to the increase in the derived ISS and NISS values when the severity of a single injury is increased from an AIS score of 4 to 5.

### 6.6.3 Reliability

Reliability describes whether a rating measures the same thing each time. A rating is considered reliable if it has a sufficiently low degree of variation relative to the total variation of all of the measurements it is intended to measure. Interrater reliability is the degree in which two raters score a single performance similarly. Different statistics are appropriate to determine inter-rater reliability for different types of measurements.
Intra-class correlation coefficient

In Paper IV, reliability was estimated using intra-class correlation coefficient (ICC) statistics and the corresponding 95% CI based on a two-way mixed model (ICC \([3,1]\)) in which the cases were random effects and the raters were fixed effects\(^{207}\). The absolute agreement measure was applied in the analysis. ICC statistics give a number on a scale from 0 to 1, where 0 indicates agreement no better than that expected by chance and 1 indicates perfect agreement\(^{208, 210}\). The ICC can be defined as the ratio of two variances:

\[
\text{ICC } [3,1] = \frac{\text{Variance due to rated subjects (cases)}}{\text{(Variance due to subjects + Variance due to raters + Residual variance)}}
\]

Hierarchical agglomerative clustering

In Paper IV, the inter-rater ICC values were further analyzed using a hierarchical agglomerative cluster (HAC) analysis with complete linkage and ‘1-ICC’ as a distance measure for the accompanying dendrogram\(^{211}\). HAC is a class of multivariate statistical methods for partitioning values into optimally similar groups based on measures of similarity\(^{211}\). A dendrogram is a graphical, tree-like representation of the distance between all pairs of ICCs in which each step of clustering is represented as a fusion of two tree branches into a single branch. The branches represent clusters that are obtained upon each step of hierarchical clustering. In this analysis, similar elements are linked near the bottom of the graph, whereas dissimilar elements are linked near the top\(^{211}\).

Linear regression

In Paper IV, univariate linear regression analysis was performed to evaluate the relationships between raters’ coding experiences, which were independent variables, and the ICC values, which were dependent variables. We were unable to perform more detailed multivariate linear regression analyses because only ten
raters were included in the study. Statistical significance was assumed for a P-value <0.05.

6.7 Ethical approvals

The Regional Committee for Medical and Health Research Ethics (REK) of Southeast Norway (ref. no. S-08499d, 2008/14588) and the Norwegian Social Sciences Data Services (ref. no. 18900) approved the project in Paper III under the premise that the exported datasets were anonymous. Written, signed consent from the centers stating that their participation and data sharing was in compliance with their own specific institutional and/or national legal frameworks and data protection requirements was a prerequisite for participation.

The Data Privacy Ombudsman for research at OUH considered the projects in Paper IV exempt from license requirements, and the REK was informed but considered the project outside their mandate for approval (ref. no. 1.2009.1139, 2009/345-1). The project was also approved by OUH-U (project no. 1435), and only raters who gave written informed consent were allowed to participate. As described above, all direct and indirect identifiable patient information was replaced by fictitious data, and hospital system identifiers, dates, and times were deleted.
7. RESULTS

7.1 Paper I

All registries, except Odense University Hospital based their inclusions on the ISS, with ISS cut-offs ranging from 9 to 16, and half of the registries included patients who were admitted to the hospital upon trauma team activation regardless of ISS score. There was also no uniform approach to inclusion of pre-hospital deaths.

The median (range) number of registered data variables was 147 (71 to 257), but only 16 variables could be considered common. Four variables had the same response categories but did not have the exact same definitions. One registry recorded comorbidity. Three systems recorded ‘survival at end of acute care’, two recorded ‘30 day survival’, and one recorded both. Only one registry had a precise data definition catalog. Each registry collected variables that enabled TRISS methodology probability of survival analyses, but some registries recorded continuous RTS values and others recorded RTS categorical values.

This discrepancy indicates that the data could not be compared with acceptable validity nor be incorporated into a multicenter trauma registry study of epidemiological significance. It was recommended that selection criteria, core data variables, and data variable definitions was unified by a consensus-based method.

7.2 Paper II

The Utstein Trauma Template was revised. NISS >15 was chosen to define severe trauma and as the single inclusion criterion, regardless of trauma team activation or intensive care unit admission. Five exclusion criteria and 36 core data variables, with four subsidiary variables, were agreed upon, and it was recommended that excluded patient groups would be considered separately.

The dataset was divided into the following groups: 1) ‘Predictive model variables’, 2) ‘Systems characteristics descriptors’, and 3) ‘Process mapping variables’. Group 1
included predictors reported to be prognostic\textsuperscript{212}. Group 2 enabled the description of differences in the care of individual patients across trauma systems and the evaluation of system structure effects on patient outcome. Group 3 variables were intended to measure trauma care process efficiency for individual patients, which could potentially be used as trauma care process indicators.

It was recommended that the newest available AIS edition be used at all times for injury severity classification. However, it was assumed that different AIS editions did not represent a major problem for outcome evaluation, when compared to coder variability including factors such as different levels of AIS training. For short-term survival, the expert panel suggested ‘30-day mortality’.

Variables that were considered but were not included as core variables are presented in Appendix A.

\textbf{7.3 Paper III}

The feasibility of collecting most of the core data was demonstrated across several registries and countries. Of the 36 Utstein Trauma Template variables, 13 (36\%) were collected by all participating centers, and 34 variables were collected by >80\% of the centers. Patient-level data completeness >80\% was achieved for 28 variables, and completeness >90\% was achieved for 20 variables. Three variables (age, gender, and AIS codes) were documented in all patients, and the variables ‘Time Until Normal Arterial Base Excess’, ‘Arterial Base Excess’, and ‘Pre-Hospital Respiratory Rate’ were the least complete. Time variable difficulties and a lack of outcome variable uniformity and injury scoring systems were found. Eleven (46\%) centers applied survival outcome definitions that varied from the template. Seventeen (71\%) centers used the recommended AIS version, but the number of AIS codes recorded per patient differed, and two centers only reported the seventh digit AIS severity code because they used ICD-9-CM-to-AIS mapping. The lack of uniformity
in outcome variables and injury scoring systems across international trauma institutions indicate the need for better standardization. The current results may serve as a stepping-stone towards establishing a European trauma registry.

7.4 Paper IV
Ten of the 19 invited raters answered a questionnaire, and each scored 50 cases. They assigned 3561 AIS codes, with a median (range) of 352 (275 to 459), while 382 AIS codes were assigned in the reference standard. Of the AIS codes, 2189 (61.5%) agreed with the reference standard. A total of 471 (13.3%) contained errors but were still considered relevant and were therefore included in the analyses. The raters recorded 392 injury codes that did not exist according to the reference standard, overlooked 1187 (31.1 %) injuries, and double-coded 509 injuries (e.g., coded more than one skin injury).

The raters found 15 injuries that were not found by the reference panel, six severity levels were changed in the reference standard, and nine injuries were removed.

The analyses of LoAs between all raters and between each rater and the reference standard for ISS showed that the narrowest LoA ranged from -8.12 to 10.48 and that the widest ranged from -35.98 to 35.22. Rater 1 disagreed markedly with all other raters and the reference standard, while raters 3 and 4 agreed the most with each other and the reference standard.

The analysis of LoAs between all raters and between the raters and the reference standard for NISS showed that the narrowest LoA ranged from -15.29 to 17.97 and that the widest ranged from -42.25 to 33.21. As for NISS, Rater 1 was the least in agreement with all other raters and the reference standard. Rater 3 agreed the most with the reference, and raters 4 and 6 agreed the most with each other. All LoAs, both for ISS and NISS, were wider than the predefined, clinically acceptable range of ±9.
The joint (range) ICC for all raters against the reference standard was 0.51 (0.29 to 0.86) for ISS and 0.51 (0.27 to 0.78) for NISS. The joint (range) ICC for inter-rater reliability was 0.49 (0.19 to 0.85) for ISS, and 0.49 (0.16 to 0.82) for NISS. Hierarchical agglomerative clustering of inter-rater-estimated ISS ICC values revealed two subgroups, with raters 1 and 9 being in the least agreement with the other raters. These two clusters showed relatively little agreement with one another. The highest degree of agreement was observed between raters 3 and 4, followed by raters 6 and 8. Inter-rater NISS ICC clustering also showed two subgroups, confirming that raters 1 and 9 were the least in agreement with the rest. The highest degree of agreement was observed between raters 8 and 10. The most narrow ISS LoAs for within-panel inter-rater agreement ranged from -10.69 to 9.65, and the widest ranged from -16.87 to 13.63. The narrowest NISS LoAs ranged from -13.44 to 16.24, and the widest ranged from -25.75 to 18.83. All of these ranges were wider than the predefined, clinically acceptable LoA range of ±9. Furthermore, the joint (range) ICC for inter-rater reliability was 0.72 (0.64 to 0.79) for ISS and 0.68 (0.64 to 0.77) for NISS. The relatively low levels of agreement and reliability of injury severity scoring indicate that summative injury scoring using the AIS system is subject to large inter-rater variability and thus must be interpreted with caution.
8. DISCUSSION AND CONCLUSIONS

There has been no European consensus on trauma data collection and reporting, and no pan-European registry exists for auditing outcomes and the quality of care of severely injured patients24, 59. This PhD project aimed to initiate a consensus process for a common European core dataset and to evaluate the quality of a few selected variables by methodological studies. While U.S. models have been widely adopted worldwide, the dataset should not only include data and prognostic models that have been developed in other continents or in other health care settings.

The focus of the PhD projects was to develop a feasible, reliable, and valid common core dataset. Papers I and II differ from the other papers. While Paper I describes the similarities between data variables and selection criteria in Scandinavian trauma registries, the main objective of Paper II was achieved by an expert panel. None of these projects included patient data. Paper III focused on the probability that the template could be applied to patient-level data, and Paper IV investigated the template from the perspective of the coder/rater by testing rater-variability in scoring patient-level data. Paper III had less of a pan-European focus, but the inclusion of more countries in the project indicated a gradually rising interest in a common international template.

8.1 Consensus method concerns

Consensus was reached on a set of core data for a potential European trauma registry. However, some limitations should be mentioned. It should be noted that 13 out of the 19 panelists were Scandinavian; therefore, some major European countries did not participate. This skewed involvement carries a significant risk. Involving a larger number of clinicians from other countries may have yielded a
different list of variables and could have improved the scientific value and the
likelihood of a wider adoption of the template.
The scientific value of methods to achieve consensus, such as Delphi Surveys and
NGT processes, has been questioned\textsuperscript{213,214}, and no method is considered to be the
“gold standard”\textsuperscript{215}. Nevertheless, consensus methods continue to be useful for
agreement on matters where hard evidence is difficult to obtain. A potential
problem with group meetings, such as those in Paper II, may be that verbally
skilled panelists monopolize the group with arguments over wording\textsuperscript{198} and that
charismatic members and more experienced registry researchers might take control
of the consensus process to defend their own preferences. Furthermore, we cannot
rule out that members of larger European registries were given more “talk time”.
The project group tried to avoid these issues by making the first two stages
anonymous and by completing the ranking of data variables prior to the first
meeting. However, it was judged important that experts with extensive knowledge
within the field of trauma registry development and research were allowed open
and free sharing of their viewpoints in the group meeting.
The results of consensus studies derive their credibility, in part, from the
composition of the consensus panel\textsuperscript{215}. The fact that more than 500 trauma-receiving
hospitals and registries across Europe have implemented the Utstein Trauma
Template data variables to date may indicate that the selection of panelists and the
conclusions that were reached were appropriate. It may also mean that the
subsequent implementation job was given priority and, to a certain point, was
successful.

8.2 Core dataset concerns
It is extremely important that data variables in trauma comparison projects are
collected in a uniform manner. For this reason, each variable and response category
of the Template was specifically and thoroughly designed and defined. Furthermore, a set of selection criteria was developed to avoid selection bias.

The Template is far from perfect. At the time of development, many trauma registries had difficulties obtaining data on patients where more than one unit and institution were involved, particularly when patients were transferred between hospitals. Therefore, the expert panel based their consensus on the premise that the core dataset was intended to cover the main hospital where a patient was treated. Thus, the variables that were agreed upon represent data for admissions to the first hospital within 24 hours after injury. This decision makes total trauma system comparisons difficult. Although some countries may have developed national inclusive trauma systems and registries, others have limited themselves to hospital-based registries. The exclusion of patients that are transferred >24 hours after injury may strongly influence trauma care comparisons when hospitals with large proportions of transferred patients are included. This issue must be addressed in a dataset update.

The panel did recommend that all trauma registries develop methods to track patients through the trauma system (i.e., from the scene of injury to discharge from the hospital or rehabilitation center) and that both the primary (local) trauma hospital and the referral trauma center record the same set of core data variables.

It is important to realize that no indicator is 100% valid in all situations and for all trauma systems; therefore, careful judgment is needed during data collection and research. The identification of the “correct” variables and the development of common definitions is a complex and on-going process. Other scientists would probably have recommended different or additional variables, but we hope that our study is an important step towards a population-based European Trauma Registry with a common data template.

While the Template will not be adopted by every hospital that collects trauma data immediately, the 500 hospitals/registries that presently use it should be able to demonstrate its benefits over time. We hope that through further standardization
processes, more institutions may join in designing a feasible, international trauma core dataset. To increase the feasibility and implementation (i.e., acceptance) of the template in Europe, the Utstein group plans to increase the number of countries that are represented in future dataset updates. Although a uniform regional, national, or international dataset may provide data that are easily collected from the vast majority of trauma patients and may answer basic epidemiological and outcome questions, it is important to keep in mind that trauma outcome is multifactorial. Hence, such a dataset can never provide the whole truth and solution.

8.3 Feasibility concerns
Feasibility may be tested in several ways. A strength of this thesis was that the method of testing feasibility improved from Paper I to Paper III. The method in Paper I may be criticized for not being scientific enough, but this limitation was emphasized during the planning and execution of Paper III.

Paper I add to the literature on data collection variability between trauma registries. A study by Mann et al. in 2006 from the U.S. showed that many states maintained a centralized registry but that the requirements for data submission varied significantly. Similar to Paper I, inclusion and exclusion criteria varied\textsuperscript{23}. Paper I indicated that Scandinavian trauma registries had few common core data and data definitions. All registries enabled TRISS calculations, but the inclusion criteria varied too much to ensure a valid comparison. One limitations of this study was that the selection of large Scandinavian trauma registries may have introduced a bias, but very few trauma registries actually existed at the start of the project. Therefore, it was judged important to prioritize the leading trauma-receiving hospitals and registries. Paper III indicated that the majority of variables in the developed Template were feasible, although some registries experienced difficulties collecting some of the
essential variables. It is interesting that ‘30-day survival’, a variable that has been recommended by the United Nations216, 217, implemented in the European Commission’s CARE-database218, and recommended by the Organization for Economic Cooperation and Development219, is not widely implemented in European and international trauma care research. This variable is probably easier to use in Scandinavian countries, where National Population Registries exist220 and enables hospitals to check patient outcome (dead or alive) using national, unique patient identifiers. This is not the case in other parts of Europe; therefore, both ‘in-hospital survival’ and ‘30-day survival’ should probably be included in the Utstein Template until all trauma centers can check 30-day survival status more easily. It is also worth noting that although all participants in Paper III used the AIS system, 30% of the participating institutions did not use the recommended AIS version. Several recent studies have identified differences between the AIS 1998 and 2005/2008 dictionaries in the number of patients that were classified as ‘major trauma’109-111, 221, 222, illustrating that data that is collected using different AIS versions cannot be directly compared. Palmer et al. demonstrated that when using the mapping system that is available in the AIS dictionary, 13% of the AIS 1998 injuries studied could not be mapped to an AIS 2008 code111. When comparing the outcomes and ISS/NISS values between trauma registries, such differences could affect the discrimination between severely and less severely injured patients across registries. Ongoing AIS dictionary updates and a lack of contemporary, precise, and easily applicable mapping systems between dictionaries are likely to hamper trauma volume and performance comparisons across and within institutions.

In Paper III, the number of AIS codes that were recorded per patient per institution differed. Incomplete AIS-coding of all injuries may prevent a detailed understanding of the patterns and extent of injuries that are sustained per patient. Furthermore, two centers only reported the seventh digit of the AIS severity code. Without the first six digits, there was no information on injured body regions and anatomic structures and no opportunity to identify selected groups of injuries.
The process that was applied for identifying potential centers was subjective and not standardized; therefore, the participating centers may have been more likely to comply with the Utstein Template or better able to collect and report the requested data. The ten centers that did not respond to the invitations and the four that agreed to participate but never submitted the requested materials were not contacted again. Thus, we cannot exclude the possibility that these centers found collection of the dataset too difficult or time consuming. Participation may have increased if more time had been spent on the follow-up of non-responding centers. Future Utstein Template feasibility analyses may focus more on evaluating its applicability. It would be of interest to determine how the variables could be combined to compare systems, processes, and outcomes across trauma registries and systems. This may reveal further shortcomings of and possibilities with the template.

8.4 Reliability concerns

In Paper IV, we initially aimed to evaluate data variable reliability based on a European sample of trauma registry coders. However, because such a process was too time and resource demanding for a PhD thesis, we included only Norwegian raters. While we only tested a few variables in this project, these variables represent crucial explanatory factors (predictors) that are considered important for case-mix adjustments in prognostic research and trauma care benchmarking.

Paper IV showed that in our current setting, the use of the AIS-methodology as a basis for injury severity summation would not have adequate benchmarking precision. It may be that targeted training, more comprehensive AIS-courses, injury coding consensus processes, more time to code at each hospital, regular recertification, and properly designed databases may increase the accuracy, agreement, and reliability of this methodology to an acceptable level. Overall, these findings should instigate quality improvement processes. Because a more adequate
and precise injury-scoring alternative is not currently available, the trauma community may use these results to improve the scoring accuracy and precision of the current scoring methods. Doing so is especially important when introducing a national inclusive trauma registry system.

We designed Paper IV to resemble the daily work of a trauma coder. The raters were not instructed to code a list of pre-defined injuries or pre-existing diseases, which would only test reliability in a “laboratory-like” setting, but were instructed to identify and score data from real patient records without assistance. This study did not test the direct reliability of the AIS scale; rather, it evaluated the accuracy of injury identification and the reliability of ISS and NISS values. Therefore, we were not able to judge the actual reliability of AIS 2008. Further studies are required to test the reliability of the AIS scale.

The inclusion of only patients with a NISS >15 may have introduced a sample selection bias. The reliability might have been different for less injured patients, who are more often treated and scored in lower-level trauma centers. Therefore, future studies should test the reliability of AIS scoring in patients with mild to moderate injuries.

Nine invited hospitals declined, did not respond to the invitation to participate in the study, or withdrew from the study. A lack of response from the invitees may have introduced a selection bias, although we have no reason to assume that the characteristics of the non-respondents differed greatly from those of the respondents. Participation may have been increased by increasing the financial compensation or by utilizing the database of the future Norwegian National Trauma Registry.

Future studies may include an intra-rater reliability test and a new rater-against-reference standard reliability test. Further studies may also be designed to evaluate how differences in scoring between raters affect the utility of the ISS and NISS in outcome prediction models for trauma patients. Additionally, more studies are needed to evaluate the reliability of the other Utstein variables.
8.5 Validity concerns

Validity can be defined as the degree to which a variable or score measures what it is intended to measure. There are four main types of validity: face validity, content validity, construct validity, and criterion validity. Within the time frame of this PhD thesis, validity was only briefly touched upon because it took more than one year after developing the template for data collections to start. We have touched upon face validity in Paper II and criterion validity in Paper IV, but much work remains for testing the validity of the template variables.

There are several methods to evaluate validity. Criterion validity can be tested by estimating the correlation of a scale or variable with some other measure that is used and accepted in the field, which is ideally a “gold standard.” Another method would be to evaluate the sensitivity, specificity, and predictive value of a scale or variable.

8.6 Developing a core dataset and ethical considerations

Benchmarking performance and outcome analysis will probably have an increasing role in health care. As health care professionals, we owe it to the public to organize our trauma systems to be safe, effective, equitable, and patient-centered according to the principles of best practice. It is time to leave the focus on quality assessment and improvement of the “islands of experience” (the trauma centers) and focus on the entire trauma system. Therefore, quality assessment systems and data collection tools that can obtain data from all services and institutions involved in a trauma system are needed. A common dataset is a requirement for benchmarking and for efforts to improve the quality of care of severely injured patients across Europe. Using insufficient data definitions and selection bias, i.e., insufficient methodology, may lead to erroneous conclusions on the quality of care.
8.7 A European trauma registry

Through the development of the Utstein Template, further initiatives towards a pan-European Trauma Registry have been made. During the project period, we established the European Trauma Registry Network (ETRN)\textsuperscript{225}, which aims to improve the management of severely injured patients in Europe. Extensive and comprehensive research in this wide field is often beyond the means of individual European hospitals and can best be accomplished through multi-disciplinary, multi-national efforts of scientists and clinicians.

One aim of the ETRN has been to develop a European, minimum mandatory, common dataset and data dictionary to serve as a basis for a data collection process for a formal European trauma registry. The Utstein Trauma Template\textsuperscript{24, 55} which has been suggested as the European core dataset has been implemented by several European trauma-receiving hospitals\textsuperscript{225}, and further registries have announced their support.

Several steps that remain in the development and establishment of a European trauma registry are briefly listed in Appendix B.

8.8 Challenges met during the projects

One of the most demanding challenges of the consensus project in Paper II was to understand and take into account all of the differences between health and trauma care systems. In addition, the development of the Utstein Template Data Dictionary\textsuperscript{55} that was initiated after the panel consensus was demanding, detail-oriented, and time-consuming.

In Paper III, data collection worked quite well, although data were sampled from institutions from different countries, most of which used different trauma registries. It was still considered important to perform extensive data checks, and much time was spent removing data inconsistencies. It was a challenge that the
Template allowed the use of blank fields in some cases of missing data, which resulted in extensive correspondence with some institutions. The work with the reliability project was complicated and demanding. The process of submitting applications to the privacy ombudsman and the regional ethics committee, the identification, development, and anonymization of 50 cases, and the development of web-based data collection tools, invitations, and case distributions took more than one year. Thereafter, one and a half years were spent on data collection, the development of a reference standard, rating evaluations, and data analyses.

Lessons that were learned during all of these processes may be of help to the ETRN, the Norwegian National Trauma Registry, and the Norwegian Air Ambulance Foundation.

8.9 Some conclusions and further perspectives

The core dataset should undergo further development to allow for the tracking of patients through an inclusive trauma system. We anticipate that this will be a focus during future development of the Utstein Template. The same core data should be recorded in primary, secondary, and tertiary trauma centers and should be supplemented by a hospital-specific dataset, a trauma system-specific dataset, and a nation-specific dataset. These data should be “fixed”, not patient-specific. A hospital-specific dataset should provide information such as the number of hospital beds and the trauma center level. A trauma system-specific dataset should provide a picture of the population and area that is covered by the trauma system and should include information on how the trauma system is organized. A nation-specific dataset should contain information such as how the health care system is organized, socio-economic indicators, educational level indicators, gross domestic product, and indicators of health care expenditure (Appendix D). Furthermore, a set of defined performance, quality, and efficiency indicators for quality
improvement processes that are based on the template data should be formulated and agreed upon.
9. REFERENCES


40. National Trauma Data Bank. Reference Manual: Background, caveats, and resources: American College of Surgeons Committee on Trauma; 2005
85. Millham FH, LaMorte WW. Factors associated with mortality in trauma: re-evaluation of the TRISS method using the National Trauma Data Bank. *J Trauma.* 2004;56:1090-1096.


103. Association for the Advancement of Automotive Medicine. Available at: http://aaam.org/.


APPENDIX A

Several data variables were discussed for inclusion in the Utstein Trauma template (Paper II). The variables that were not judged as absolutely needed were the following:

Predictive model variables:
1. In-hospital place of death
2. Core temperature on admission
3. SaO₂ on hospital admission
4. Survival at end of acute care

System characteristics descriptors:
1. Time from dispatch to hospital arrival
2. Patient underwent an interventional radiological procedure
3. Did the hospital have an interventional radiologist?
4. Time from arrival to transfer
5. Admitted at a critical care unit
6. When rehabilitation started
7. Seniority of the trauma team leader
8. Ward of admittance
9. Sepsis
10. SIRS
11. Facilities
12. Time of first chest x-ray
13. Type of first bedside/portable diagnostics (e.g., chest x-ray/ultrasound)
14. Type of non-portable diagnostics (e.g., CT-scan, angiography)
Process mapping variables:

1. Time to adequate oxygenation delivery
2. Chest tube time
3. Intubation time
4. Time of accident/injury
5. Patients transferred for specialized care
6. Date and time of admission to primary hospital
7. Date and time of discharge from primary hospital
8. Date and time of transfer to secondary hospital
9. Length of hypoperfusion
   a. Base excess
   b. Dialysis
10. Evacuation of intracranial hematoma >25 ml (or equivalent)
11. Time until surgery for bleeding (time to and type of initial intervention)
12. Time until stabilization of pelvic fractures
13. Type of stabilization of pelvic fractures
14. Time until revascularization of pulseless limb
15. Time until control of decontamination in bowel surgery
16. Incidence and duration of hypothermia
17. Early CT scan
18. Glucose control
19. Time from alarm to dispatch
20. Readmission to the ICU
21. Time of first bedside/portable diagnostics
22. Time of procedure for stopping massive bleeding (damage control surgery)
APPENDIX B

Future steps in establishing a European trauma registry

1. Further improve and expand the common core (minimum) dataset
2. Implement the common dataset in European trauma centers and registries
3. Finalize the website
4. Establish a collaboration with Europe WHO
5. Establish an international advisory board for the ETRN
6. Seek funding
7. Establish an administration and secretary
8. Decide upon the location of the database
9. Develop an application tool for a European trauma registry
10. Studies on system differences, processes, performance, and outcome
11. Structured description of national emergency care systems in Europe
12. Describe the burden of severe injury in each European country
13. Develop/agree on a European anatomic injury severity classification system
14. Initiate a European trauma outcome study
15. Develop a European prognostic model(s)
16. Define a set of European indicators for quality improvement
APPENDIX C

The Abbreviated Injury Scale – AIS

The AIS classification system is an anatomically based seven-digit injury scoring system. The first six digits of a code identify the injured body region, the type of anatomic structure and the specific anatomic structure. The seventh digit refers to an injury severity scale that classifies the severity of individual injuries by body regions on a six-point ordinal scale.

<table>
<thead>
<tr>
<th>AIS code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Minor</td>
</tr>
<tr>
<td>2</td>
<td>Moderate</td>
</tr>
<tr>
<td>3</td>
<td>Serious, not life-threatening</td>
</tr>
<tr>
<td>4</td>
<td>Severe, life threatening, survival probable</td>
</tr>
<tr>
<td>5</td>
<td>Critical, survival uncertain</td>
</tr>
<tr>
<td>6</td>
<td>Maximal (currently untreatable)</td>
</tr>
<tr>
<td>9</td>
<td>For cases of missing information about injury severity</td>
</tr>
</tbody>
</table>

There are nine AIS Body Regions: Head; Face; Neck; Thorax; Abdomen; Spine; Upper Extremity; Lower Extremity; and External and other.

The Injury Severity Score – ISS

The AIS is the foundation for the ISS. ISS quantifies an overall severity of the anatomic injuries and is calculated by summing the squares of the highest AIS severity codes in each of the three most severely injured ISS Body Regions. Injury Severity Scores range from 1 to 75. If an injury is assigned an AIS value of 6, the ISS score is automatically assigned to 75, regardless of any other injuries that may also be found.

\[
ISS = (AIS)^2 + (AIS)^2 + (AIS)^2
\]
Only six ISS Body Regions to which injuries can be assigned exist, although the AIS dictionary is divided into 9 anatomical chapters (AIS Body Regions).

The ISS Body Regions:
1. Head or neck
2. Face
3. Chest
4. Abdominal or pelvic contents
5. Extremities or pelvic girdle
6. External

The New Injury Severity Score – NISS
The NISS, which is a revised version of the ISS, is calculated by summing the squares of the three most severe AIS injuries regardless of body region. The NISS also ranges from 1 to 75.

Patient example
Male, 32 years old, motorcycle collided against a large truck at high velocity. The patient was comatose at the scene of the injury and had injuries in three body regions.

<table>
<thead>
<tr>
<th>ISS body region</th>
<th>Injury descriptor</th>
<th>AIS code</th>
<th>Highest AIS</th>
<th>AIS² (ISS)</th>
<th>AIS² (NISS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head or Neck</td>
<td>Diffuse, axonal injury involving corpus callosum</td>
<td>140627.5</td>
<td>5</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Subdural hematoma; tiny Brain edema; absent ventricles</td>
<td>140651.3</td>
<td>3</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>140674.5</td>
<td>5</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Face</td>
<td>Pneumothorax NFS</td>
<td>442202.2</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Abdominal or pelvic content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extremities or pelvic girdle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External</td>
<td>Skin laceration; minor</td>
<td>210602.1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sum</td>
<td>ISS = 30</td>
<td>NISS = 59</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D

Reflections on the development of a common dataset

1. Clear purpose(s) of the registry
An important issue in the development of a common dataset and registry is to define a clear purpose. This will also affect the selected data variables.
- What measurements are desired?
- Why?
- Which types of analyses are wanted?
- What will the analyses be used for?

2. Data definition catalog (Data dictionary)
To prevent the risk of misinterpretation, an effort should be made to define each data variable as unambiguously as possible. Thus, before data variables are operational, it must be clarified how they are defined, measured, and documented. To meet this requirement, a dataset or dictionary should contain information about the following:
- Data variable name
- Abbreviated data variable name
- Definition of the data variable
- Type of data
- Data variable categories or values
- Source of data information
- Coding guidance

3. Selection criteria
To avoid selection criteria bias, the inclusion and exclusion criteria must be clearly defined before data collection begins and before comparisons of care across registries are performed. Performance analyses should ideally highlight differences
in patient care and system structure, but not differences in patient selection and definitions. If a local, regional or national registry is a part of a larger international registry (e.g., European trauma registry), the inclusion criteria for the international registry must be minimally met, i.e., have the same core criteria, and, at the same time, have the flexibility to allow expansion of the inclusion criteria at local levels to be able to satisfy local program needs. The inclusion and exclusion criteria are also important to identify and describe the population that is being studied.

4. Factors that may affect variation trauma care quality across systems

It is important to document factors that may affect variation in the quality of care across systems\textsuperscript{12}. These factors can be divided into patient factors, system factors, process factors, outcome factors, and regional/national factors.

4.1 Patient factors

The analyses should be adjusted for patient factors/characteristics. If a hospital receives patients with more severe injuries than other units, an adjustment must be made. Patient factors are also important to describe the included population. Some examples of patient factors are:

- Severity of injury
- Age, gender
- Vital signs in the pre-hospital setting and on admission
- Co-morbidities (e.g., drug and alcohol abuse, medications, or pre-existing diseases)
- Socioeconomic conditions (e.g., education and income)

4.2 System factors

The system dimension describes the resources and organization that are available (e.g., economy, personnel resources, physical frames, medical equipment, or the existence of clinical guidelines). It may also indicate differences in the organization
of care in different systems (e.g., the systems factors described in the Utstein Trauma Template).

Some examples of system factors are:

- Number of nurses with trauma competence at the departments
- Number of ICU beds
- Availability of specialists (e.g., neurosurgery or cardio-thoracic surgery)
- Does the hospital have a helipad in close proximity to the hospital?
- Highest level of pre-hospital care provided to a patient
- Type of transportation used to transport the patient to the hospital
- Hospital size

4.3 Process factors

The process dimension describes the activities that the health care service performs for the patient.

Some examples of process factors are:

- Total pre-hospital time
- Type of diagnostic examination(s) for each diagnosis
- Time from hospital arrival to emergency intervention
- Length of stay for each condition
- Which treatment the patients with diagnosis X received

4.4 Outcome factors

For the health care services, one of the most important measures of quality is what happened to the patient (i.e., the outcome). Outcome variables are measures of the results of an injury, e.g., short-term survival and functional outcome. As in a research study, a uniform registry must have outcome variables that are identical across centers, e.g., a clear uniform definition of the time of measurement.

Some examples of process factors are:

- 30-day lethality
• Six-month lethality
• Functional Independence Measure (FIM)
• Disability Rating Scale (DRS)
• SF-36 (e.g., functional health and well-being from the patient’s point of view)
• Complications / side effects / adverse events

4.5 Regional/national factors

A regional/national-specific dataset could contain indicators of the organization of the health care system, life expectancy and wellbeing, socio-economic indicators, educational level indicators, and indicators of health care expenditure.

Some examples of key regional/national indicators are:

• Gross domestic product
• Incidence of injury
• Per capita expenditure on health and average earnings
• Age-specific and/or injury-specific mortality rates\textsuperscript{227}
FEASIBILITY OF COMPARING CORE DATA FROM EXISTING TRAUMA REGISTRIES IN SCANDINAVIA. REACHING FOR A SCANDINAVIAN MAJOR TRAUMA OUTCOME STUDY (MTOS)

K. G. Ringdal1, 2, H. M. Lossius1, 2, 3, The SCANTEM ad hoc group on Scandinavian MTOS and Trauma Registry

1 Acute Care Medicine Research Network, Department of Health Studies, University of Stavanger, Stavanger, Norway
2 Department of Research and Development, The Norwegian Air Ambulance Foundation, Drøbak, Norway
3 Scandinavian Networking Group for Trauma and Emergency Management (SCANTEM)

The SCANTEM ad hoc group on Scandinavian MTOS and Trauma Registry: E. F. Christensen (Denmark), H. Grønborg (Denmark), L. Handolin (Finland), A. Leppäniemi (Finland), H. M. Lossius (Norway), O. Reise (Norway), M. Schultz-Larsen (Denmark), P. Örtvenwall (Sweden)

ABSTRACT

Background: The organisation of trauma care in Scandinavia has several similarities, including trauma registries, but so far there are limited amount of research on efficiency and outcome. Data and results from trauma outcome studies like the US MTOS are not fully applicable to the Scandinavian trauma population.

Aims: To reveal the feasibility of using data from existing trauma registries of major hospitals in Scandinavia, for a minimal common dataset, in a joint, prospective Scandinavian MTOS.

Material and Methods: We collected data points, data point definitions, and inclusion/exclusion criteria, from the major trauma registries of the Swedish trauma registry standard, three university hospitals in Denmark, one university hospital in Finland, and the Norwegian National Trauma Registry. The collected material was compared to reveal common data points, inclusion criteria, and the compatibility of data point definitions.

Results: The median number of data points was 147 (range 71–257; interquartile range = 90–205). Most registries lacked precise data definition catalogues. Only 16 data points could be considered as common, of which just a few were core trauma data. Four data points had the same data category options but were not considered having the same data point definitions. The inclusion criteria were not uniform.

Conclusions: Trauma registries in Scandinavia have few common core data and data point definitions. There were data points for calculating the Trauma and Injury Severity Score (TRISS) but the inclusion criteria varied too much to ensure a valid comparison.

Correspondence:
Kjetil G. Ringdal, M.D.
The Norwegian Air Ambulance Foundation
P.O. Box 94
N - 1441 Drøbak
Norway
Email: kjetil.ringdal@snla.no
A consensus process for a joint trauma core data set will be initiated by the Scandinavian Networking Group for Trauma and Emergency Management (SCANTEM) to increase research on trauma efficiency and outcome.

Key words: Trauma Registry; Scandinavia; MTOS; SCANTEM; data definition; core data; data uniformity; data comparison; inclusion criteria

INTRODUCTION

The organisation of trauma care in Scandinavia has several similarities including trauma registries and trauma care, but so far there has been a limited amount of joint research on efficiency and outcome. The Scandinavian Networking Group for Trauma and Emergency Management (SCANTEM) (the former Nordic Trauma Forum) was founded in 2004 (1). One main mission of SCANTEM is to promote, initiate, and coordinate research on trauma, and emergency care in Scandinavia (Denmark, Finland, Norway, and Sweden). Through SCANTEM, there has been an increasing awareness of the importance of collaborative reviews and research in trauma care from these countries.

Data and results from trauma outcome studies like the United States Major Trauma Outcome Study (US MTOS) (2) is not fully applicable to the Scandinavian trauma population. We know that the mechanism of injury for the Scandinavian trauma population differs compared to the US MTOS trauma population (3). Studies conducted in other countries, based on the Trauma and Injury Severity Score (TRISS) (4), have shown difficulties in using this methodology in populations where mechanisms of injury, co-morbidity, demographics, and local factors do not resemble the US MTOS population (3, 5–8). In contrast to the hospitals entering data into the US MTOS base which tend to receive the majority of their patients directly from the scene of injury (5, 6), many of the larger trauma hospitals in Scandinavia receive a high number of patients referred from rural community hospitals (9). In the US National Trauma Data Bank (NTDB), the traditional TRISS has shown limited ability to predict survival after trauma (10). We know that the TRISS methodology has major limitations in many subgroups of patients, especially in severe trauma (11), where the TRISS model tends to over predict survival, with the greatest effect among patients with severe injuries (12). In patients with multiple injuries in a single body region, the Injury Severity Score (ISS) (13) thus considers only one of the injuries within that region (14).

Trauma registries have been the basis for research and quality assessment of trauma management, and have informed doctors and politicians about methods for optimal care of injured patients (15). The American College of Surgeons requires that all Level I, II, III and IV trauma centres that are certified have a trauma registry (16, 17).

In Europe, the United Kingdom Trauma Audit and Research Network (UK TARN) (18) along with the Trauma Registry of the German Society for Trauma Surgery (DGU-TR) (19), and the Italian National Registry of Major Injuries (RIT) (20), represents the largest trauma registries. There has been a move towards developing a European Trauma Audit and Research Network (EuroTARN) (21), and a core data set with inclusion and exclusion criteria has been defined. The EuroTARN is still in an early phase of development, and up to date no consensus on the detail and extent of the dataset between countries has been reached (22). The first report from EuroTARN concludes that it is possible to collect data from established trauma registries and that initial analysis reveals significant international variations (23).

No joint Scandinavian trauma registry has been initiated, though several individual hospitals in Scandinavia have established local trauma registries. Sweden has developed a trauma registry standard, KVIITRA (24), and Norway has reached consensus on inclusion/exclusion criteria and core data in the Norwegian National Trauma Registry (NNTR). Individual hospitals in Finland and Denmark participate in the TARN in addition to maintaining registries of their own. As a consequence SCANTEM wants to develop a joint core data set for a Scandinavian Trauma Registry (TR), with the aim of initiating a future prospective Scandinavian MTOS for cross border comparative studies, taking special Scandinavian factors into consideration. Such an initiative would represent a major step towards developing regional norms for research and organisation of trauma care in Scandinavia. A Scandinavian TR could be based on already collected data from existing local trauma registries, but whether data from the different registries in Scandinavia are comparable, or if the inclusion and exclusion criteria are uniform remains uncertain. The Scandinavian TR will support and guide the national registries, but will also contribute to a future European trauma registry like the EuroTARN.

The objective of our study was to reveal the feasibility of using data from existing trauma registries from major trauma hospital in Scandinavia for a minimal common dataset in a joint prospective Scandinavian MTOS. We also wanted to describe the differences and similarities of the different registries in Scandinavia with an emphasis on structure, inclusion and exclusion criteria, data point definitions, and common data points.

MATERIAL AND METHODS

We collected inclusion and exclusion criteria, core data points, and data definitions from the KVIITRA, the trauma registries of three university hospitals in Denmark, one
Feasibility of comparing core data from existing trauma registries in Scandinavia

university hospital in Finland, and the NNTR (Table 1). All examined registries provided datasheets with their respective core data points. Unstructured interviews and e-mail correspondence was used to deepen the information on the datasheets. The comparison of data was based on the material received from the participating centres at the initiation of the study in 2005. Based on the received documents we compared the inclusion and exclusion criteria, pointed out common core data from patients involved in major trauma, and revealed to what degree the data point definitions were uniform.

A data point was defined as an unambiguous term without possible misinterpretation (25). The data points definition standard should include ‘Data point name’, ‘Exact definition of data point’, ‘Field format’, ‘Variable category’, and ‘Source of information’. Guidelines for recommended data point definitions was found in the data dictionary of the US NTDDB (26), TARN (27), NNTR, KVITTRA, the Injury Surveillance Guidelines (ISG) of the World Health Organization (WHO) (28), and the Utstein Style of reporting data from major trauma (29). These guidelines set a standard for how a precise data point definition can be described. The Utstein style has been the main international recommendation for uniform reporting of data following major trauma, and is a concept for uniform reporting of data, and suggests a set of important core data points from the pre-hospital phase, early in-hospital management, as well as co-morbidity and outcome data.

We compared the core data from the participating trauma registries according to the following categories of the EuroTARN core data set: General information/demographics, Injury mechanism, Pre-hospital care, Transport, Hospital care, On-going care, Outcomes, Scoring.

RESULTS

INCLUSION AND EXCLUSION CRITERIA

All registries except Odense trauma registry based their inclusion of patients on the ISS. The ISS cut off for inclusion varied from 9 to 16 (Table 2). NNTR, Odense, and Århus included all patients admitted with trauma team activation (TTA). There were no uniform approaches to the inclusion of patients who were dead at scene (Table 2).

DATA POINTS

The large number of different data points and data point definitions, made a comparison of data from the participating trauma registries difficult. We detected a diverging precision of the data point definitions in the six trauma core data sets. The median number of data points was 147 (range 71–257; inter-quartile range = 90–205). Trauma registries with the lowest number of data points were Århus (Denmark) and NNTR (Norway), and trauma registries with the highest number of data points were Töölö (Finland) and Rigshospitalet (Denmark). When comparing the registries, we found 16 common data points (Table 3). Among these 16 data points, four data points had similar data category options, but were not considered having common definitions. These were ‘Date of trauma’, ‘Date of discharge’, ‘Intubation at scene’, ‘In-hospital intubation’. The data points Glasgow Coma Scale (GCS) (30), Systolic Blood Pressure, and Respiratory Rate were defined according to either the exact continuous values or according to the Revised Trauma Score (RTS) values (31), and the data point definitions were hence not considered fully common.

We have no information on which version of the Abbreviated Injury Scale (AIS) (32) the participating registries used.

DATA DEFINITION CATALOGUES

There were varying degrees of the precision of data point definitions. The NNTR had defined and categorised all their core data points precisely, with detailed instructions for data collection and source of information. We found many different field description and number of categories (e.g. primary/secondary, blunt/penetrating, text field for injury description etc.).

COMMENTS ON INDIVIDUAL REGISTRIES

The KVITTRA trauma registry had 112 data points. They had definitions for their data points. In the KVITTRA we found 12 data points for pre-hospital data, but few data points for information on Intensive Care Unit (ICU) treatments and observations.

The Norwegian National Trauma Registry had 83 data points. This registry had precise and defined data points. The data were categorised and contained data on source of information, as well as instructions for correct coding. In the NNTR there were 8 data fields for pre-hospital data, but no data on CT-scanning or operative treatments. The NNTR had no ICU data points.

The Odense trauma registry had 181 data points. There were a text field for detailed description of trauma, and data fields for registering use of airbag, use of helmet, means of transportation, and position in car. In the Odense registry there were 10 ‘check off’ fields for interventions done pre-hospital. There were no data points for “on scene” physiological parameters, and no data on ‘Time of arrival at scene’. The ICU data points were ‘Date of admission’ and ‘Date of discharge’. The datasheets from the Odense trauma registry showed no data fields for registering ISS, New Injury Severity Score (NISS) (33), or TRISS, even though they do calculate the values.

<table>
<thead>
<tr>
<th>Country</th>
<th>Trauma registry/hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>Odense University Hospital (Odense)</td>
</tr>
<tr>
<td></td>
<td>Rigshospitalet University Hospital (Copenhagen)</td>
</tr>
<tr>
<td>Norway</td>
<td>Norwegian National Trauma Registry (NNTR)</td>
</tr>
<tr>
<td></td>
<td>– Stavanger University Hospital (Stavanger)</td>
</tr>
<tr>
<td></td>
<td>– Ullevål University Hospital (Oslo)</td>
</tr>
<tr>
<td>Sweden</td>
<td>KVITTRA (Swedish trauma registry standard)</td>
</tr>
<tr>
<td>Finland</td>
<td>Töölö Hospital (Helsinki University Hospital)</td>
</tr>
</tbody>
</table>
The Rigshospitalet trauma registry had 257 data points. They used the data definitions from UK TARN (27). The registry recorded the number of doctors involved, as well as operation procedures, grade and specialty of doctors involved. Rigshospitalet was the only hospital that included data on co-morbidity in their trauma registry. The registry had 24 data fields for pre-hospital data, but few data fields for information on ICU treatments and observations. The registry calculates the NISS and TRISS but we found no data point name for that in the provided datasheets.

The Töölö trauma registry had 213 data points. They used the data point definitions from the UK TARN. Data points for patient demographics like nationality, residence, and postal code were present in the Töölö registry. There were 24 pre-hospital data fields in the Töölö registry. The Töölö registry was the only registry with data point collection on interventions done en route. The registry is currently implementing a new trauma registry with a new data dictionary.

The Århus trauma registry had 71 data points that were partially defined according to the data point definition standard. The registry had a text field for detailed description of trauma. This registry included data on the departments involved in the treatment of the patients. In the registry there were 7 pre-hospital data fields. These pre-hospital data were taken from ambulance charts. There were no data points for “on scene” treatment. The Århus registry had a minimum
TABLE 3

Common data points from the examined trauma registries/hospitals.
The data points were grouped according to the EuroTARN data categories.

<table>
<thead>
<tr>
<th>DATA POINT GROUPS</th>
<th>COMMON DATA POINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>General information/demographics</td>
<td>Age (e.g. classification)</td>
</tr>
<tr>
<td></td>
<td>Date of trauma</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
</tr>
<tr>
<td>Injury mechanism</td>
<td>Injury type</td>
</tr>
<tr>
<td>Pre-hospital care</td>
<td>Intubation at scene</td>
</tr>
<tr>
<td>Transport</td>
<td>Date of ED arrival</td>
</tr>
<tr>
<td></td>
<td>Transportation</td>
</tr>
<tr>
<td>Hospital care/ED</td>
<td>Respiratory rate</td>
</tr>
<tr>
<td></td>
<td>Systolic blood pressure</td>
</tr>
<tr>
<td></td>
<td>In-hospital intubation</td>
</tr>
<tr>
<td></td>
<td>X-ray</td>
</tr>
<tr>
<td>On going care/ICU</td>
<td>Total number of ICU days</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Date of discharge</td>
</tr>
<tr>
<td>Scoring</td>
<td>AIS</td>
</tr>
<tr>
<td></td>
<td>In-hospital GCS</td>
</tr>
<tr>
<td></td>
<td>In-hospital RTS</td>
</tr>
</tbody>
</table>

Note: ED = Emergency Department; ICU = Intensive Care Unit; AIS = Abbreviated Injury Score; GCS = Glasgow Coma Score; RTS = Revised Trauma Score; EuroTARN = European Trauma Audit & Research Network.

of ED data points, and there were no ICU data other than data for ‘Date of admission’ and ‘Date of discharge’.

DISCUSSION

There is a wide variability of content in trauma registries in Scandinavia. The participating registries show diverging core data sets, different number of core data, and differing data point definitions. We found only 16 common core data points. Some of the demographic data, and only two clinical data had common categorical options.

There were data points for recording outcome, but some registries had data points for ‘30 day mortality’, while others had data points for ‘dead/alive’ at discharge. We believe that this is neither a uniform data point nor a uniform data point definition because the outcome status consequently can be measured at different point in time.

The comparison of core data and inclusion/exclusion criteria revealed that each trauma registry collect data as a basis for the TRISS methodology even though some registries use exact values for the RTS and others use the coded value (0–4) according to the RTS. However, the inclusion and exclusion criteria were not uniform, and by using the stated inclusion criteria (Table 3) we can not be sure that all patients with an ISS > 15 will be included in the TRISS evaluation. Hence, the lack of homogeneity in inclusion criteria will not make the trauma populations comparable.

A variation in the edition of the AIS coding dictionary used will most probably not be of major importance for a comparison of trauma registries (34). Some of the registries included a high number of data points in their trauma registry; others set their data collection at a minimum. It has been shown that the data elements collected by trauma registries internationally varies considerably (17), and studies on the validity of local databases have reported both over- and underestimation (35, 36). We therefore recommend that trauma registries minimize the number of core data collected due to the potential of incomplete datasets (9). The temptation to collect core data for interest rather than usefulness should probably be resisted (37). If the goal is epidemiology as well, data could be collected as recommended by the Utstein style, in core data (i.e. always obtained) and in optional data (i.e. obtained under specific circumstances) (29).

Mann et al. recommend several important steps for allowing individual registry data to be compared among countries and regions (15). Amongst others, they recommend developing a uniform list of data points with standard data point definitions that could represent a minimal core data set, with individual registries collecting other data points of interest. Because of the wide variability among registry-inclusion criteria, Mann et al. recommend that a least-common-denominator approach may be necessary (15). The guidelines laid down by the Utstein style (29) for classifying, defining and coding the data point set, should be used as a basis for a uniform set of data variable definitions. The data dictionary of a Scandinavian TR could also be based on the recently developed Data Dictionary from the US National Trauma Data Bank (NTDB) (26).

We found that only one of the participating trauma registries had data fields for describing and registering patient co-morbidity. According to Jurkovich et al. co-morbidity factors known to significantly affect outcome should be included in all registries (38). Bouamra et al. argues that the absence of reliable data points for pre-existing diseases in the TARN makes it impossible to include co-morbidity in models used for benchmarking (39). A recent work from Norway reports pre-injury American Society of Anesthesiologist (ASA) physical status as an independent predictor of mortality after trauma (40), but the use and implementation of co-morbidity data in outcome studies needs further elucidation.

In our opinion trauma data from the examined trauma registries cannot be compared with acceptable validity; nor can they, at the time being, be incorporated into a prospective Scandinavian multicentre trauma registry study of epidemiological significance.

Each of the trauma registries examined seem to have been created for slightly different purposes and perspectives, suggesting that the various registries were established to address varying objectives, and implying that the system improvement goals of the trauma registries vary across Scandinavia. A future Scandinavian system will need enough core uniformity to make comparison. We recommend a unifica-
CONCLUSIONS

The major trauma registries in Scandinavia have nearly identical core data, but there is data for TRIS and the Krone criteria, which will have an influence on the definition of trauma core data. The Euro-TRIS project is initiated by the Scandinavia Trauma and Injury Research Group and is supported by the International Trauma Audit and Research Network (ITAN) and the European Trauma Society (ETS).

Received: August 8, 2007
Accepted: October 9, 2007
The Utstein template for uniform reporting of data following major trauma: A joint revision by SCANTEM, TARN, DGU-TR and RITG

Kjetil G Ringdal*1,2, Timothy J Coats3, Rolf Lefering4, Stefano Di Bartolomeo5, Petter Andreas Steen2, Olav Røise6, Lauri Handolin7, Hans Morten Lossius1 and Utstein TCD expert panel

Address: 1Department of Research, Norwegian Air Ambulance Foundation, Drøbak, Norway, 2Faculty of Medicine, Faculty Division Ullevål University Hospital, University of Oslo, Norway, 3Academic Unit of Emergency Medicine, Leicester University, UK, 4Institute for Research in Operative Medicine, University of Witten/Herdecke, Cologne-Merheim Medical Centre, Cologne, Germany, 5Unit of Hygiene and Epidemiology, DPMSC, School of Medicine, University of Udine, Italy, 6Orthopaedic Centre, Ullevål University Hospital, Oslo, Norway and 7Department of Orthopaedics and Traumatology, Helsinki University Central Hospital, Finland

Email: Kjetil G Ringdal* - kjetil.ringdal@snla.no; Timothy J Coats - t.coats@virgin.net; Rolf Lefering - rolf.lefering@ifom-uni-wh.de; Stefano Di Bartolomeo - stefano.dibartolomeo@med.uniud.it; Petter Andreas Steen - p.a.steen@medisin.uio.no; Olav Røise - olav.roise@medisin.uio.no; Lauri Handolin - lauri.handolin@pp.inet.fi; Hans Morten Lossius - hans.morten.lossius@snla.no; Utstein TCD expert panel - kjetil.ringdal@snla.no

* Corresponding author

Abstract

Background: In 1999, an Utstein Template for Uniform Reporting of Data following Major Trauma was published. Few papers have since been published based on that template, reflecting a lack of international consensus on its feasibility and use. The aim of the present revision was to further develop the Utstein Template, particularly with a major reduction in the number of core data variables and the addition of more precise definitions of data variables. In addition, we wanted to define a set of inclusion and exclusion criteria that will facilitate uniform comparison of trauma cases.

Methods: Over a ten-month period, selected experts from major European trauma registries and organisations carried out an Utstein consensus process based on a modified nominal group technique.

Results: The expert panel concluded that a New Injury Severity Score > 15 should be used as a single inclusion criterion, and five exclusion criteria were also selected. Thirty-five precisely defined core data variables were agreed upon, with further division into core data for Predictive models, System Characteristic Descriptors and for Process Mapping.

Conclusion: Through a structured consensus process, the Utstein Template for Uniform Reporting of Data following Major Trauma has been revised. This revision will enhance national and international comparisons of trauma systems, and will form the basis for improved prediction models in trauma care.
Background

The Utstein template for uniform reporting of data following major trauma

To permit data collection and statistics on major trauma care, in 1999 a working group from the International Trauma Anaesthesia and Critical Care Society (ITACCS) published a recommendation for the Utstein Template for Uniform Reporting of Data following Major Trauma [1]. The template extracted data for the pre-hospital phase, early in-hospital management, and for co-morbidity and outcome. In accordance with the previous Utstein templates, it was commended that data were to be classified as ‘Core’ (essential) or ‘Optional’ (supplemental). Despite the intention of facilitating studies to improve the understanding of trauma and trauma care, only a few papers have been published based on the template [2,3]. This indicates a need for further development, and in particular, a major reduction in the large number (92) of core data variables [1], as well as the addition of more precise definitions of these variables [3].

Trauma registries

Due to the practical difficulties with performing randomised controlled trials in severe trauma cases, valid scientific evidence is often lacking. Systematic prospective registry-based data collection for documenting trauma care is performed by several local, regional and national trauma registries. However, such registries cannot replace randomised clinical trials, but allow for exploration of relationships present in the collected data. The primary aims of these trauma registries are to enable comparative analyses of trauma care and outcome to provide quality improvement and optimal care of the injured patients [4]. The development of a European trauma registry may provide population-based comprehensive data on trauma incidence, epidemiology and trends. Further, it may enable development of regional outcome prediction models (taking special European factors into consideration) and thus set baseline norms for future trauma outcome studies. In Europe, there has been some reluctance to share data definitions and coding formats vary significantly between registries, and also because patient selection is not comparable [5,28]. Further efforts to establish uniform and standardised inclusion and exclusion criteria, as well as a minimum list of core data variables with precise definitions, are essential [3,5]. In addition, consistent methods of injury scoring need to be agreed upon [4,29-31]. To address this need for a European consensus, SCANTEM, TARN, DGU-TR and the Italian National Registry of Major Injuries (RITG) [32] carried out a consensus process, concluding with symposia in May and December 2007 at the Utstein Abbey [33], Norway. Selected experts met with the aim of further developing the Utstein Template for Uniform Reporting of Data following Major Trauma. At that time, they defined inclusion and exclusion criteria, and a minimum core dataset with precise

TRISS methodology

Over the last two decades, the Trauma and Injury Severity Score (TRISS) method [7,8], with coefficients for prediction of outcome has been the most commonly used method for comparison of outcome in trauma patients. The TRISS coefficients were originally derived from the United States Major Trauma Outcome Study (US MTOS) [9,10] but more recently the coefficients have been updated based on patient cases from the National Trauma Data Bank [11]. However, the TRISS method has some limitations, and it has been criticised by many authors [7,12-21]. Among other things, the TRISS model requires scoring the Revised Trauma Score (RTS) [22] components (Glasgow Coma Scale [GCS] [23], respiratory rate [RR] and systolic blood pressure [SBP]) on admission in the emergency department (ED), and does not take into account co-morbidity. Despite its limitations, TRISS continues to be the most accepted and widely-used tool for comparing trauma outcome in North America and in some parts of Europe.

Comparing and benchmarking European trauma care

In Europe, the UK Trauma Audit and Research Network (TARN) [24], along with the Trauma Registry of the German Society of Trauma Surgery (DGU-TR) [25], represent the largest trauma registries. There has also been a move towards developing a European Trauma Audit and Research Network (EuroTARN) [26], and a core dataset with inclusion and exclusion criteria has been created. Nevertheless, to date, no consensus has been reached between countries on the details and extent of the dataset. A first report from EuroTARN concluded that it is possible to collect data from established trauma registries, and the initial analysis revealed significant international variation [5]. As a continuation of this effort, a European project has been initiated by the DGU-TR, UK TARN and the Scandinavian Networking Group for Trauma and Emergency Management (SCANTEM) [27], for developing a joint European Core Dataset (EuroCoreD) for a future European Trauma Registry.
definitions. In addition, the aim of the revised template was to develop a standard for comparison of trauma data that was compatible with the large trauma registries in Europe, also adhering to EuroTARN. The template was intended to support the establishment of a European Trauma Registry, promote further development of a European model for outcome prediction and allow European and international trauma auditing and benchmarking.

Methods
This revision of the Utstein template is based on a nominal group technique (NGT) process [34,35] modified to fit the purpose. For participation in the NGT process, a European expert panel was selected.

The expert panel
The expert panel was comprised of those individuals who were central to developing and managing the largest European trauma registries; the panel included clinicians, database managers and epidemiology experts.

Data variable definition
A data variable should be unambiguously defined (with no misinterpretations) and reasonably simple to register. To meet this requirement, a data variable dictionary should contain information on 'data point number,' 'data point name,' 'descriptive field name,' 'type of data,' 'data point category/value,' 'definition of data point,' 'source of data information' and 'coding guidance.' We based recommended guidelines for data variable definitions on existing trauma registry databases, the Utstein Template for Uniform Reporting of Data following Major Trauma [1], the US National Trauma Data Standard (NTDS) [36] and the Injury Surveillance Guidelines from the World Health Organization (WHO) [37].

Core data variables
A registry should differentiate between data variables that absolutely need to be collected (core data) and the type of additional data that may be desirable (optional data) [1,37]. The current revision focuses on core data that are considered to be essential for documentation and reporting. We divided the core data into three groups ('Predictive Model,' 'System Characteristic Descriptors' and 'Process Mapping Variables') based on the role of the data variable in a registry.

Predictive model
The predictive model is composed of patient and injury severity variables that are considered to be important for outcome prediction. Predictive models are not deterministic; rather, they provide the probability of an outcome for a given patient [38]. Complex models, such as Abbreviated Injury Scale (AIS) [39] derivatives and the RTS, are often used to create such predictive models [38]. Experience from the German and UK trauma registries suggests that there may be better data variables to include in a predictive model than those traditionally used in the TRISS methodology [24,40-42].

System characteristic descriptors
Data variables in the System Characteristic Descriptor group describe trauma systems. Within Europe, there are large differences in philosophies and structures of trauma care systems, and these data variables should indicate key differences between systems and permit comparisons of the effect of system structure on outcomes.

Process mapping variables
Process mapping variables are intended to describe trauma care at an individual trauma centre (e.g., what happens to a patient after a major trauma); these are used for documentation of the patient journey, care process and care activities.

Specific premises
At present, many trauma registries have difficulty in obtaining data for patients from all involved hospitals when patients are transferred between them; therefore, the expert panel based their consensus on the premise that the core dataset was intended to cover the main hospital where a patient is treated. However, the expert group recommended that all trauma registries develop methods to track patients through the trauma system and that both the primary (local) trauma hospital and the referral trauma hospital record the same set of core data variables. The introduction of a core outcome data variable will secure that the overall effect of the entire trauma system can be measured, even if part of the patient’s treatment course is not recorded in detail.

The nominal group technique
The modified NGT process consisted of four steps. First, each expert was supplied with necessary background documents (Table 1), and asked to return (by e-mail) proposals for inclusion and exclusion criteria, as well as a maximum of 30 core data variables in a prioritised order. This first proposal was summarised and structured by the coordinators (KGR, HML), and the collated results were redistributed in the second step for comments and re-prioritisation. The third step consisted of two consensus meetings in which members of the expert panel discussed their views in a structured way and then made conclusions. In the fourth step, the panellists were able to comment on the conclusions by e-mail. To complete the process, a letter of consent was signed by all experts.

Results
The expert panel concluded that a New Injury Severity Score [43] (NISS) > 15 should be used as a single inclu-
sion criterion (Table 2). Five exclusion criteria were listed (Table 2), and a total of 35 core data variables (23 in the predictive model group, eight system descriptors and four process mapping variables) were agreed upon (Tables 3, 4 and 5).

**Discussion on inclusion/exclusion criteria and core data variables**

**Inclusion criteria**

NISS is a modification of the Injury Severity Score (ISS) method [43]. ISS is calculated by summing the squares of the highest AIS severity codes in each of the three most severely injured ISS body regions [44]. Hence, ISS will ignore all but the most severe injury in a body region, and often fails to consider worse injuries in other regions of the body [43]. In contrast, NISS is defined as the sum of the square of the three most severe AIS injuries regardless of body region [43]. Several authors have argued for replacing ISS with NISS [43,45-49]. Osler et al. considered NISS to be easier to calculate and more predictive of survival than the ISS method [43], and a recent study by Lavoie et al. confirmed their findings [46]. NISS will be equal to or greater than ISS for any given patient, and it appears to be a more accurate method for rating severely injured patients [49,50]; specifically, this is true for patients with multiple head injuries [46]. The increased number of included patients by choosing NISS > 15 instead of ISS > 15 should be seen as an increase in 'sensitivity' without a loss of 'specificity' of an ideal definition of major trauma. An effort should be made to secure that all patients with a NISS > 15 are included, regardless of whether or not the trauma team was activated prior to or upon the patient's arrival at the hospital, and whether or not the patient was admitted to an intensive care unit.

**Exclusion criteria**

Using NISS > 15 as a single inclusion criterion will include some patients that are at high risk of confounding data analysis. To remove such patients from the analysis, a set of exclusion criteria was defined. The expert panel recommended excluding first hospital admissions more than 24 hours after the injury (e.g., prolonged search and rescue missions), patients declared dead before hospital arrival, or those with no signs of life (pupillary response, spontaneous ventilation, presence of carotid pulse, measurable or palpable blood pressure, extremity movement, or cardiac electrical activity) [51] upon hospital arrival and those having no response to hospital resuscitation. In addition, it was recommended that asphyxias, drowning and burns should be excluded (Table 2).

Pre-hospital deaths should be excluded for practical reasons, since in some countries patients declared dead in the pre-hospital setting are transported directly to the morgue; whereas in other countries, they are admitted to hospital. All patients who arrive in the ED with spontane-

---

**Table 1: Attachments sent to the expert panel prior to the Utstein 2007 meeting.**

<table>
<thead>
<tr>
<th>No.</th>
<th>Document name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dick et al. Recommendations for uniform reporting of data following major trauma – the Utstein style [1].</td>
</tr>
<tr>
<td>2</td>
<td>Conclusions from the Utstein symposium on ‘Improving Trauma Systems and the Role of Trauma Registries’.</td>
</tr>
<tr>
<td>3</td>
<td>Inclusion and exclusion criteria and data points from the European Trauma Audit &amp; Research Network.</td>
</tr>
<tr>
<td>4</td>
<td>The Swedish Trauma Registry Standard (KVITTRA), Data Dictionary.</td>
</tr>
<tr>
<td>5</td>
<td>The Norwegian National Trauma Registry, Data Dictionary.</td>
</tr>
<tr>
<td>6</td>
<td>American College of Surgeons, National Trauma Data Bank; National Trauma Data Standard, Data Dictionary v. 1.2 [36].</td>
</tr>
<tr>
<td>7</td>
<td>ICD-10, Chapter XX. External causes of morbidity and mortality [61].</td>
</tr>
</tbody>
</table>

**Table 2: Inclusion and exclusion criteria.**

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>NISS &gt; 15.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclusion criteria</td>
<td>First hospital admission more than 24 hours after injury. Patients declared dead before hospital arrival, or with no signs of life on hospital arrival and no response to hospital resuscitation. Asphyxia. Drowning. Burn patients should be excluded if the burn represents the predominant injury, or if the patient is treated in a specialised burn unit.</td>
</tr>
</tbody>
</table>

NISS: New Injury Severity Score [43]. Signs of life: Pupillary response, spontaneous ventilation, presence of carotid pulse, measurable or palpable blood pressure, extremity movement, or cardiac electrical activity [51].
Table 3: Predictive model variables.

<table>
<thead>
<tr>
<th>Data variable no.</th>
<th>Data variable name</th>
<th>Type of data</th>
<th>Data variable categories or values</th>
<th>Definition of data variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Age</td>
<td>Continuous</td>
<td>Number</td>
<td>The patient’s age at the time of injury.</td>
</tr>
<tr>
<td>2</td>
<td>Gender</td>
<td>Nominal</td>
<td>1 = Female 2 = Male 3 = Unknown</td>
<td>The patient’s gender.</td>
</tr>
<tr>
<td>3</td>
<td>Dominating Type of Injury</td>
<td>Nominal</td>
<td>1 = Blunt 2 = Penetrating 3 = Unknown</td>
<td>Indication of the type of injury produced by the trauma.</td>
</tr>
<tr>
<td>4</td>
<td>Mechanism of Injury</td>
<td>Nominal</td>
<td>1 = Traffic: motor vehicle injury (car, pickup truck, van, heavy transport vehicle, bus) 2 = Traffic: motorcycle injury 3 = Traffic: bicycle injury 4 = Traffic: pedestrian 5 = Traffic: other (ship, airplane, railway train) 6 = Shot by handgun, shotgun, rifle, other firearm of any dimension 7 = Stabbed by knife, sword, dagger, other pointed or sharp object 8 = Struck or hit by blunt object (tree, tree branch, bar, stone, human body part, metal, other) 9 = Low energy fall (fall at the same level) 10 = High energy fall (fall from a higher level) 11 = Other 12 = Unknown</td>
<td>The mechanism (or external factor) that caused the injury event. The cut-off level for a fall should be defined as the person’s height.</td>
</tr>
<tr>
<td>5</td>
<td>Intention of injury</td>
<td>Nominal</td>
<td>1 = Accident (unintentional) 2 = Self-inflicted (suspected suicide, incomplete suicide attempt, or injury attempt) 3 = Assault (suspected) 4 = Other 5 = Unknown</td>
<td>Information about the role of human intent in the occurrence of an injury, primarily determined by the incident and not by the resulting injury.</td>
</tr>
<tr>
<td>6</td>
<td>Pre-injury ASA-PS Classification System</td>
<td>Ordinal</td>
<td>1 = A normal healthy patient 2 = A patient with mild systemic disease 3 = A patient with severe systemic disease 4 = A patient with severe systemic disease that is a constant threat to life 5 = A moribund patient who is not expected to survive without the operation 6 = A declared brain-dead patient whose organs are being removed for donor purposes 7 = Unknown</td>
<td>The pre-injury co-morbidity existing before the incident. Derangements resulting from the injury should not be considered.</td>
</tr>
<tr>
<td>7</td>
<td>Pre-hospital cardiac arrest</td>
<td>Nominal</td>
<td>1 = No 2 = Yes 3 = Unknown</td>
<td>Did the patient suffer an injury-related pre-hospital cardiac arrest?</td>
</tr>
</tbody>
</table>
Table 3: Predictive model variables. (Continued)

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable Description</th>
<th>Type</th>
<th>Value Range</th>
<th>Outcome Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Glasgow Coma Scale (GCS) upon arrival of EMS personnel at scene</td>
<td>Ordinal</td>
<td>Number</td>
<td>First recorded pre-interventional GCS upon arrival at scene of medical personnel trained to assess.</td>
</tr>
<tr>
<td>9</td>
<td>GCS motor component upon arrival of EMS personnel at scene</td>
<td>Ordinal</td>
<td>Number</td>
<td>First recorded pre-interventional GCS motor component upon arrival at scene of medical personnel trained to assess.</td>
</tr>
<tr>
<td>10</td>
<td>GCS upon arrival in ED/hospital</td>
<td>Ordinal</td>
<td>Number</td>
<td>First recorded GCS upon arrival in the ED/hospital.</td>
</tr>
<tr>
<td>11</td>
<td>GCS motor component upon arrival in ED/hospital</td>
<td>Ordinal</td>
<td>Number</td>
<td>Fist recorded GCS motor component upon arrival in the ED/hospital.</td>
</tr>
<tr>
<td>12a</td>
<td>Systolic Blood Pressure (SBP) upon arrival of EMS personnel at scene</td>
<td>Continuous</td>
<td>Number</td>
<td>First recorded SBP upon arrival at scene of medical personnel trained to assess.</td>
</tr>
<tr>
<td>12b</td>
<td>SBP – clinical category – upon arrival of EMS personnel at scene</td>
<td>Ordinal</td>
<td>RTS 4 = &gt;89 (&quot;good radial pulse&quot;)</td>
<td>First recorded SBP upon arrival at scene of medical person trained to assess.</td>
</tr>
<tr>
<td>13a</td>
<td>SBP upon arrival in ED/hospital</td>
<td>Continuous</td>
<td>Number</td>
<td>First recorded SBP upon arrival in the ED/hospital.</td>
</tr>
<tr>
<td>13b</td>
<td>SBP – clinical category – upon arrival in ED/hospital</td>
<td>Ordinal</td>
<td>RTS 4 = &gt;89 (&quot;good radial pulse&quot;)</td>
<td>First recorded SBP upon arrival in the ED/hospital.</td>
</tr>
<tr>
<td>14a</td>
<td>Respiratory Rate (RR) upon arrival of EMS personnel at scene</td>
<td>Continuous</td>
<td>Number</td>
<td>First recorded RR upon arrival at scene of medical personnel trained to assess.</td>
</tr>
<tr>
<td>14b</td>
<td>RR – clinical category – upon arrival of EMS personnel at scene</td>
<td>Ordinal</td>
<td>RTS 4 = 10–29 (&quot;normal&quot;)</td>
<td>First recorded RR upon arrival at scene of medical personnel trained to assess.</td>
</tr>
<tr>
<td>15a</td>
<td>RR upon arrival in ED/hospital</td>
<td>Continuous</td>
<td>Number</td>
<td>First recorded RR upon arrival in the ED/hospital.</td>
</tr>
<tr>
<td>15b</td>
<td>RR – clinical category – upon arrival in ED/hospital</td>
<td>Ordinal</td>
<td>RTS 4 = 10–29 (&quot;normal&quot;)</td>
<td>First recorded RR on arrival in the ED/hospital.</td>
</tr>
</tbody>
</table>
Table 3: Predictive model variables. (Continued)

<table>
<thead>
<tr>
<th></th>
<th>Arterial Base Excess</th>
<th>Continuous</th>
<th>Number</th>
<th>First measured arterial base excess after arrival in the hospital.</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Coagulation: INR</td>
<td>Continuous</td>
<td>Number</td>
<td>Use the first measured INR within the first hour after hospital arrival.</td>
</tr>
<tr>
<td>17</td>
<td>Number of days on ventilator</td>
<td>Continuous</td>
<td>Number</td>
<td>The total number of patient days spent on a mechanical ventilator (including all episodes). Record in full day increments with any partial day listed as a full day.</td>
</tr>
<tr>
<td>18</td>
<td>Length of stay in main hospital treating the patient</td>
<td>Continuous</td>
<td>Number</td>
<td>Calculate 'Date of discharge' minus 'Date of admission' from the reporting hospital.</td>
</tr>
<tr>
<td>19</td>
<td>Discharge destination</td>
<td>Nominal</td>
<td></td>
<td>The patient's destination after end of acute care in the main hospital treating the patient. CCU = critical care unit.</td>
</tr>
<tr>
<td>20</td>
<td>Glasgow Outcome Scale – at discharge from main hospital</td>
<td>Ordinal</td>
<td></td>
<td>Glasgow Outcome Scale score at discharge from main hospital.</td>
</tr>
<tr>
<td>21</td>
<td>Survival status</td>
<td>Nominal</td>
<td></td>
<td>Alive or dead 30 days after injury.</td>
</tr>
<tr>
<td>22</td>
<td>Abbreviated Injury Scale (AIS)</td>
<td>Ordinal</td>
<td>Number</td>
<td>The AIS severity codes that reflect the patient's injuries. All injuries should be listed, even duplicated codes (e.g., bilateral femoral fractures, multiple spine fractures). The edition of the AIS coding dictionary should be indexed; AIS 2005 is recommended.</td>
</tr>
</tbody>
</table>

ASA-PS: American Society of Anesthesiologists Physical Status [65].
ED: Emergency Department.
EMS: Emergency Medical Services.
INR: International Normalized Ratio.
RTS: Revised Trauma Score [22].

http://www.sjtrem.com/content/16/1/7
Table 4: System characteristic descriptors.

<table>
<thead>
<tr>
<th>Data variable no.</th>
<th>Data variable name</th>
<th>Type of data</th>
<th>Data variable categories or values</th>
<th>Definition of data variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>Time from alarm to hospital arrival</td>
<td>Continuous</td>
<td>HH:MM</td>
<td>The time between when the alarm call is answered (at the emergency call centre) and when the patient arrives at the reporting hospital.</td>
</tr>
<tr>
<td>25</td>
<td>Highest level of prehospital care provider</td>
<td>Ordinal</td>
<td>1 = Level I. No Field Care</td>
<td>The highest available level of competence of the pre-hospital care providers involved in the care of the injured patient.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 = Level II. Basic Life Support</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = Level III. Advanced Life Support, No Physician Present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 = Level IV. Advanced Life Support On-Scene, Physician Field Care</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 = Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 = Unknown</td>
<td></td>
</tr>
<tr>
<td>26a</td>
<td>Pre-hospital intubation</td>
<td>Nominal</td>
<td>1 = No</td>
<td>Was the patient intubated before arrival at the hospital?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 = Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = Unknown</td>
<td></td>
</tr>
<tr>
<td>26b</td>
<td>Pre-hospital intubation</td>
<td>Nominal</td>
<td>1 = A tube in the trachea (orotracheal, nasotracheal, or surgical airway) – drug assisted</td>
<td>Type of pre-hospital intubation. Drug assisted = anaesthesia, neuromuscular blocking drugs, and deep sedation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 = A supraglottic airway adjunct that prevents speech (such as esophago-tracheal combitube, the laryngeal tube, and various kinds of laryngeal masks)) – drug assisted</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = A tube in the trachea (orotracheal, nasotracheal, or surgical airway) – not drug assisted</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 = A supraglottic airway adjunct that prevents speech (such as esophago-tracheal combitube, the laryngeal tube, and various kinds of laryngeal masks) – not drug assisted</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 = Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 = Unknown</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Type of transportation</td>
<td>Nominal</td>
<td>1 = Ground ambulance</td>
<td>Type of transportation delivering the patient to the hospital.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 = Helicopter ambulance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = Fixed-wing ambulance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 = Private/public vehicle</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 = Walk-in</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 = Police</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7 = Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 = Unknown</td>
<td></td>
</tr>
</tbody>
</table>
Ous circulation should be included, even if they have had a period of cardiac arrest before being admitted or if they die in the ED.

Asphyxia, drowning and burns are sufficiently different from blunt and penetrating injuries to require other datasets, and need to be considered separately. The UK National Burn Injury Database [52] is currently in use specifically for this purpose. Although the AIS 2005 edition has codes for asphyxia and drowning, such injuries were not included in earlier AIS editions, making comparisons across versions more difficult. In some (but not all) countries, major burn patients are sent to dedicated burn unit hospitals, thereby confounding comparisons. Burn patients should be excluded if the burn represents the predominant injury, or if a patient is treated in a specialised burn unit. In such patients, outcome is determined by factors other than those suggested in this paper. Including burn patients will not represent a sufficient number of

---

**Table 4: System characteristic descriptors.** (Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Nominal</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>Type of first key emergency intervention</td>
<td>Nominal</td>
<td>1 = Damage control thoracotomy — (any emergency or urgent thoracotomy performed for bleeding or suspected bleeding into the chest, but excluding simple thoracic tube drainage)</td>
<td>2 = Damage control laparotomy — (any emergency or urgent laparotomy performed for bleeding or suspected bleeding into the abdomen, including bleeding from the sora)</td>
<td>3 = Extraperitoneal pelvic packing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 = Limb revascularisation (Arterial injury necessitating vascular surgery or interventional radiology — including all interventions for pulseless limb, decreased perfusion and intimal arterial injuries)</td>
<td>5 = Interventional radiology (Angiographic embolisation; Stent; Stent-graft placement — excluding limb revascularisations which are classified as 4)</td>
<td>6 = Craniotomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7 = Intracranial pressure (ICP) device insertion (excluding cases were the ICP device was inserted as part of a craniotomy which are classified as 6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 29 | Activation of the trauma team | Nominal | 1 = No | 2 = Yes | 3 = Unknown |
|    | Was the patient met by an activation of the trauma team prior to or upon arrival at the hospital? |

| 30 | Inter-hospital transfer | Nominal | 1 = No | 2 = Yes – Transferred IN to the reporting hospital | 3 = Yes – Transfer OUT of the reporting hospital |
|    | Was the patient transferred from/to another hospital for acute treatment? |
|    | 4 = Yes – Transferred both IN and OUT of the reporting hospital | 5 = Unknown |

| 31 | Highest level of in-hospital care | Ordinal | 1 = Emergency Department | 2 = General Ward | 3 = Operation Theatre |
|    | The highest level of care in the main hospital. |
|    | 4 = High Dependency Unit | 5 = Critical Care Unit (definition based on nurse to patient ratio) | 6 = Unknown |
patients to report on; hence, burn-related injuries will add little power to the predictive model.

**Predictive model variables**

Age is an independent predictor of survival after trauma [53,54]. While the original TRISS model operates with only two age categories, current predictive models utilise different age groups, and we therefore recommend reporting the patient’s nominal age (continuous) at the time of injury, in years without decimals, and always rounding down. Patients under one year of age should be reported with one decimal number (e.g., six months is 0.5).

Gender is recommended as a core data, since some studies have reported no association between gender and mortality after traumatic injury [55]; whereas others have found age-specific associations between male gender and outcome [56-58].

An evaluation of type of injury (blunt versus penetrating trauma) is useful for determining which patients are candidates for surgical haemostasis [59]; whereas others have found that penetrating injuries, the predominant type of injury should be recorded [1]. The expert panel defined the dominating injury as the one with the highest AIS score. In the rare event of a patient having both blunt and penetrating trauma with the same AIS severity score, penetrating trauma is defined as the predominant injury.

The significance of the mechanism of injury (MOI) in prediction of trauma and outcome is, to a large extent, undetermined [60]. The MOI should be of value for epidemiology or subgroup analysis, and should be described in categories with reasonable prevalence rates. The International Classification of Diseases, 10th revision (ICD-10) [61], chapter XX, External causes of morbidity and mortality (V01-Y98), was initially examined for the purpose of the template; however, it was found to be too detailed, with too many injury codes. Therefore, the expert panel developed a reduced set of categories, which should make data collection easier. The set still enables the analysis of important subgroups, and since it is compatible with the ICD-10 codes, it will allow future category expansion if required.

In the ICD, most injuries can be grouped into two dimensions: intent and mechanism [62]. ‘Intention of injury’ provides information about the role of the human intent of an injury. The included list of categories is based on the ICD-10 codes, and is selected by the expert panel since it covers most injury intentions.

The presence of significant co-morbidity represents an independent predictor of mortality after trauma [1,53,63,64], and the expert panel recommends employing the American Society of Anaesthesiologists Physical Status (ASA-PS) classification system [65] for classifying the pre-injury co-morbidity status concretised by selected examples from the Norwegian Society of Anaesthesiology.

---

**Table 5: Process mapping variables.**

<table>
<thead>
<tr>
<th>Data variable no.</th>
<th>Data variable name</th>
<th>Type of data</th>
<th>Data variable categories or values</th>
<th>Definition of data variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>Time from alarm to arrival at scene</td>
<td>Continuous</td>
<td>HH:MM</td>
<td>The time from when the emergency call is answered (at the emergency call centre) until the first medical provider (at least the equivalent of EMT’s) arrives at the patient.</td>
</tr>
<tr>
<td>33</td>
<td>Time until normal arterial base excess</td>
<td>Continuous</td>
<td>HH:MM</td>
<td>The time from first measured arterial base excess at hospital admission until first measured arterial base excess within normal range. Reference range for base excess: ± 3 mmol/l.</td>
</tr>
<tr>
<td>34</td>
<td>Time to first CT scan</td>
<td>Continuous</td>
<td>HH:MM</td>
<td>The time from hospital admission until the time marked on the first CT scan image.</td>
</tr>
<tr>
<td>35</td>
<td>Time until first key emergency interventions</td>
<td>Continuous</td>
<td>HH:MM</td>
<td>The time from hospital admission until the FIRST emergency intervention. Record the time from hospital admission until the time of FIRST knife to skin is performed. Consider only the emergency interventions listed in data variable number 28.</td>
</tr>
</tbody>
</table>

CT: Computed Tomography.  
EMT: Emergency Medical Technician.
The ASA-PS classification system is easy to use and employed daily by anaesthesiologists world-wide. The pre-injury ASA-PS classification system should be used solely to categorise co-morbidity that exists before injury [6]. Skaga et al. found this classification system to be a significant predictor of mortality after trauma, also when adjusting for the variables in the traditional TRISS model [6]. Physiologic derangement resulting from the present injury is not reflected in the pre-injury score, only co-morbidity existing before the incident [6]. The pre-injury ASA-PS is not identical to the preoperative or the post-injury ASA score from anaesthesia records.

Pre-hospital cardiac arrest after trauma has dismal outcomes [67-70]; it was therefore chosen as a core data variable.

The panel recommends using raw values (continuous data) of the GCS, SBP, RR, where they are obtainable, and coded values (clinical categories) according to the RTS [22] (Table 7) in those cases missing raw values. In addition, the expert panel recommends recording the first pre-hospital, pre-interventional GCS, SBP and RR collected by a trained person, as well as the first GCS, SBP and RR upon arrival at the hospital.

For patients intubated before arrival at the hospital, the ED GCS is not attainable, and this could potentially result in large amounts of missing data [18]. For this reasons, the pre-hospital, pre-interventional GCS should be recorded, thus permitting its inclusion in future prediction models. Some studies have shown that the pre-hospital GCS score correlate well with the arrival GCS score [71,72], and that the pre-hospital GCS score is a strong predictor of outcome [73]. A change in the GCS from the field to hospital arrival is highly predictive of outcome [72,74]. In the RTS, GCS carries the greatest weight [22], and the motor component of GCS appears to contain most of the predictive power of the total GCS score [75]. This component needs further assessment, something that may be facilitated by reporting both the pre-hospital pre-interventional GCS motor component as well as the first recorded GCS motor component upon arrival in the ED/hospital.

Chan et al. found that even in the case of normal SBP upon arrival in the ED, out-of-hospital hypotension was a clinical predictor of severe injury [76]; it was also a strong predictor of the need for emergency surgery, according to Lipsky et al. [77]. Others have documented an increased risk of intra-abdominal injury in blunt trauma patients with hypotension in the hospital [78,79]. Franklin et al.

<table>
<thead>
<tr>
<th>Table 6: American Society of Anesthesiology Physical Status (ASA-PS) Classification System.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ASA-PS 1</strong></td>
</tr>
<tr>
<td>Guidelines:</td>
</tr>
<tr>
<td>Example:</td>
</tr>
<tr>
<td><strong>ASA-PS 2</strong></td>
</tr>
<tr>
<td>Guidelines:</td>
</tr>
<tr>
<td>Examples:</td>
</tr>
<tr>
<td><strong>ASA-PS 3</strong></td>
</tr>
<tr>
<td><strong>ASA-PS 4</strong></td>
</tr>
<tr>
<td>Guidelines:</td>
</tr>
<tr>
<td><strong>ASA-PS 5</strong></td>
</tr>
<tr>
<td><strong>ASA-PS 6</strong></td>
</tr>
</tbody>
</table>

The six ASA-PS headings are from the American Society of Anaesthesiologists [65]. The guidelines and examples were translated from the Norwegian edition [66] by Skaga et al. [6]. For the Utstein Template, the ASA-PS classification system should solely be used to categorise pre-injury comorbidity. Derangements resulting from the injury should not be considered.
reported 16% mortality in patients with a period of pre-hospital hypotension but with stable vital signs on ED presentation, vs. 27% mortality for patients with normal pre-hospital SBP who developed hypotension in the ED [80]. Changes in the SBP values, from the pre-hospital pre-interventional value to the SBP value on hospital arrival, could give valuable information about physiologic derangements.

Missing RR values for patients arriving in the hospital (patients intubated before arrival) is the most common cause of a lack of RTS values [81]. Collecting the pre-hospital, pre-interventional RR as well as allowing the use of coded values instead of precise values may compensate for this [18]. In addition, continued documentation of RR for later determination of its predictive power is highly recommended. Creating alternative predictive models may also necessitate the continuous collection of RR values.

Admission arterial base excess (BE) is a predictor of mortality after trauma [40,82-84]; it is also considered an indicator of haemodynamic instability, high transfusion requirement, and an indicator of metabolic and coagulatory decompensation in trauma patients [40,85]. Other results indicate that arterial BE is a predictor of intrabdominal injury [78,79], and the European guidelines for management of bleeding following major trauma have recommended arterial base deficit as a sensitive test to estimate and monitor the extent of bleeding and shock [59]. Kroezen et al. showed that replacing RTS by base deficit as a measure of physiological disturbance could predict mortality as well as RTS in the TRISS model [86]. The expert panel recommends recording the first arterial BE after arrival in the hospital.

As outcome measures, the expert panel suggest hospital LOS, discharge destination, Glasgow Outcome Scale (GOS) [91] score at hospital discharge and 30-day mortality. The GOS contains information about morbidity, and this data collection represents an increased focus on information regarding morbidity. In the original US MTOS, "end of acute care" was used as an outcome measure [9]; however, 30-day mortality is considered to be a more fixed endpoint than "end of acute care/hospital discharge" or "end of somatic care," which will vary depending on the transfer and rehabilitation policies of an individual system [92]. Using 30-day mortality is consistent with the previous Utstein recommendations [1]. Death occurring later than 30 days after injury is more likely to be caused by other conditions, such as pre-existing disease [81]. An analysis of 69,650 patient admissions from the TARN showed that 4.8% of the patients died within 93 days of admission [41]. Of these, only 9% died later than 30 days after admission; these were mainly patients with a low ISS (< 9) and aged > 65 years. In a Scandinavian trauma registry, 4.6% of the deaths occurred

Table 7: Revised Trauma Score (RTS) categories with clinical notes.

<table>
<thead>
<tr>
<th>RTS coded values</th>
<th>Respiratory Rate</th>
<th>Systolic Blood Pressure</th>
<th>Glasgow Coma Scale score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>10–29 (&quot;normal&quot;)</td>
<td>&gt;89 (&quot;good radial pulse&quot;)</td>
<td>13–15</td>
</tr>
<tr>
<td>3</td>
<td>&gt;29 (&quot;fast&quot;)</td>
<td>76–89 (&quot;weak radial pulse&quot;)</td>
<td>9–12</td>
</tr>
<tr>
<td>2</td>
<td>6–9 (&quot;slow&quot;)</td>
<td>50–75 (&quot;femoral pulse&quot;)</td>
<td>6–8</td>
</tr>
<tr>
<td>1</td>
<td>1–5 (&quot;slow&quot;)</td>
<td>1–49 (&quot;only carotid pulse&quot;)</td>
<td>4–5</td>
</tr>
<tr>
<td>0</td>
<td>0 (&quot;no respiration&quot;)</td>
<td>0 (&quot;no carotid pulse&quot;)</td>
<td>3</td>
</tr>
</tbody>
</table>

This table is based on (but not identical to) the RTS table in reference [22]. The parentheses represent clinical notes that were added by the expert panel.
later than 30 days after injury [92], whilst data from the German trauma registry (cases with NISS ≥ 16) indicate that 4.9% of the patients that died did so later than 30 days after injury [personal communication with the DGU-TR]. Survival status of all patients at a single given endpoint is needed, and 30-day mortality was chosen to represent this information.

For injury severity description, the expert panel recommends recording all AIS codes. It has previously been shown that a comparison of survival for trauma registries that use different AIS editions is potentially possible [93]. A difference in registrars and different levels of AIS training probably represents a greater problem than problems with using different AIS versions. However, registries are recommended to use the same AIS edition as a uniform way of coding; hence, the newest available version should be used at all times. At present, this is the AIS 2005 edition.

System characteristics descriptors
The expert panel considered the time between when the alarm call is answered (at the emergency dispatch centre) and when the patient arrives at the first hospital to be an important system characteristic descriptor (this variable can also be useful for mapping the entire process of pre-hospital rescue) and recommend that this interval be reported.

Recording the highest level of competence of the pre-hospital care providers was regarded by the expert panel as an important measure for describing the pre-hospital system, and for cross-border comparisons of pre-hospital trauma care and outcome. The highest available level of the providers may vary somewhat in Europe, but the revised template’s categorisation of level of provider is based on the levels proposed by McSwain [94], since these levels will encompass most pre-hospital systems.

Pre-hospital intubation is an important parameter that represents pre-hospital advanced life support (ALS) [95]. There is an ongoing debate on the use and role of ALS [96-98] measures in the out-patient management of trauma victims, and such information should be made available for possible inclusion in future survival prediction models. The original US TRISS model was derived from a dataset that excluded trauma patients who were intubated out-of-hospital [9]. However, these patients constitute a significant proportion of European trauma victims; hence, the expert panel recommends including them. In a study by Arbabi et al., early field intubation was associated with a decreased risk of fatal outcome, as compared to ED intubation [71]. In their study, intubation status was also an independent predictor of fatal outcome, after adjusting for ISS, SBP, mechanism, age and ED-GCS [71].

The type of transportation delivering the patient to the hospital is an important descriptor of the pre-hospital trauma care system, but the use of a helicopter vs. ground ambulance remains controversial [99,100], and assessment is confounded by differences in various EMS and HEMS systems. Since transportation type has been so widely debated, this is recommended as core data to be collected. The data variable does not cover what is a commonly used transport combination in some parts of Europe; ground ambulance to the local hospital and fixed-wing or rotary-wing transfer to the regional trauma centre. The data variable was developed to cover the type of transportation delivering the patient to the reporting hospital.

The expert panel recommends registering the initial key emergency intervention (EI) conducted during the hospital stay (ED, OR, critical care unit). These interventions (Table 8) represent essential emergency procedures used for the treatment and stabilisation of patients with severe injuries. Some registries will probably collect finer resolution data on interventions, but most patients will fit into one of these categories. It is recommended that the EI found in the present categories be recorded, even if there is no proof in a patient’s notes that the cause of intervention was bleeding. The term “damage control”, as used in some categories, implies that only the urgent (rather than later planned procedures) should be recorded. All limb

<table>
<thead>
<tr>
<th>No.</th>
<th>Emergency interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Damage control thoracotomy – (any emergency or urgent thoracotomy performed for bleeding or suspected bleeding into the chest, but excluding simple thoracic tube drainage)</td>
</tr>
<tr>
<td>2</td>
<td>Damage control laparotomy – (any emergency or urgent laparotomy performed for bleeding or suspected bleeding into the abdomen, including bleeding from the aorta)</td>
</tr>
<tr>
<td>3</td>
<td>Extraperitoneal pelvic packing</td>
</tr>
<tr>
<td>4</td>
<td>Limb revascularisation (Arterial injury necessitating vascular surgery or interventional radiology, including all interventions for pulseless limb, decreased perfusion and intimal arterial injuries)</td>
</tr>
<tr>
<td>5</td>
<td>Interventional radiology (Angiographic embolisation; Stent; Stent-graft placement – excluding limb revascularisations which are classified as No. 4)</td>
</tr>
<tr>
<td>6</td>
<td>Craniotomy</td>
</tr>
<tr>
<td>7</td>
<td>Intracranial pressure (ICP) device insertion (excluding cases where the ICP device was inserted as part of a craniotomy, which are classified as No. 6)</td>
</tr>
</tbody>
</table>
revascularisations should be categorised in the same group (whether surgical or radiological) since it is the fact that the patient needed revascularisation (rather than the exact method used) that is important. If radiological revascularisations are put together with all the rest of the interventional radiology, the frequency of revascularisations undertaken in a particular system will not be easily measurable. Information about the surgical revascularisations will not permit comparison of the need for revascularisation across systems. The need for revascularisation is a more interesting comparison than comparing the way in which revascularisations were achieved. Extraperitoneal pelvic packing for control of massive traumatic pelvic haemorrhage has been described by several authors [101-103], and is a separate EI subcategory. Rapid intervention is essential for patients with intracerebral haematomas that require evacuation; for patients requiring only intracranial pressure (ICP) monitoring as part of the overall intensive care management of diffuse brain injuries [104], however, this intervention would not have the same level of urgency. Nevertheless, insertion of an ICP monitoring device will provide information about neurosurgical alertness in a trauma system, and is therefore included. The suggested list of EI was developed to cover the broad majority of emergency interventions (but perhaps not all of them). For the purpose of comparative evaluation of the acute treatment process, some less common types of intervention may not be appropriate.

A formal trauma team is an essential part of an organised trauma system [104]. Activation of the trauma team is an organised initial response to a trauma [104,105] with the primary goal of securing fast and efficient treatment of severely injured patients [106-108]. Information on whether the trauma patient was met by an activation of the trauma team was recommended as vital for describing a trauma system.

The revised template recommends recording the highest level of in-hospital care for the trauma patient in the main trauma hospital (Table 4).

**Process mapping variables**

The “Chain of Survival” concept emphasises that all time-sensitive interventions must be optimised to maximise the chance of patient survival [109]. The expert panel recommends recording the time between when the emergency call is answered (at the emergency call centre) and when the first medical provider arrives at the patient. This core data represent parts of the first link in that chain, and is an important measure of the quality of the pre-hospital EMS system.

As an overall marker of the efficiency of patient treatment (including resuscitation, diagnostics and surgery), the expert group suggested to consider the time required to achieve normal arterial BE. A good evidence basis does not exist for this recommendation, however, if BE worsens after arrival, prognosis worsens as well [110]. The BE should be measured regularly after hospital arrival, and the template advises to document the initial measurement of BE immediately after hospital arrival, and in cases of abnormal values, to document the time in hours until normalisation.

The efficiency of the initial in-hospital management is assessed by the time from hospital admission until the time marked on the first CT scan images. This data represents the time required to perform key in-hospital diagnostic tests.

The time to the first key EI should be recorded; this measure represents how quickly an urgent intervention fundamental to the treatment and stabilisation of a patient’s specific injuries [105] is performed. This core data measures the efficiency of the trauma system in the initial phase. The time elapsed between injury and EI should be minimised.

**General discussion**

The present paper represents a further development of the previous Utstein Template for Uniform Reporting of Data following Major Trauma [1], and reflects a need for the creation of new prediction models that are more suitable for the type of trauma seen in Europe. Currently, both the German and UK trauma registries have developed their own predictive models, and no longer use the original TRISS system. However, uniform and standardised inclusion and exclusion criteria and a core list of data variables with precise definitions are mandatory before comparisons between trauma registries can be made. The expert panel reached a consensus on such a list of core data. This dataset does not preclude the possibility that in local trauma registries more data can legitimately be considered core information for specific purposes. The data agreed upon in the present process represent data for admissions to the first hospital within 24 hours after injury. We are aware that some countries have made further progress in developing national trauma systems and national trauma registries; whereas others only register information at local trauma hospitals. Currently, the core dataset will be difficult to use to assess the entire trauma system of some regions or countries, making direct system comparisons difficult. At this point, the expert panel recommends focusing on the hospitals of definitive treatment, and registering patient transfers (as a separate data variable) should allow for this when data are analysed. The inclusion of 30-day mortality in predictive models allows for the assessment of more of the total system performance. However, excluding pre-hospital deaths and the details of
management along the chain that precede the main treatment hospital represents a major limitation in total system assessment. The core dataset should undergo further development to allow for routine tracking of the patient through a trauma system consisting of more than one hospital. We anticipate that future development of the Utstein Template will focus more on the trauma system as European trauma registries develop, making it possible to follow transferred patients. In future trauma systems, the same core data variables should be recorded in primary, secondary and tertiary trauma centres.

The template classifies fall-related injuries as low or high (category options of the MOI data variable in Table 4) with a person’s height as the cut-off value. Various definitions of falls are used internationally [61,111-115]; they range from different categories of height, different cut-off levels for height, to differing units of height measurement (feet vs. meters). This is an area in which there is a lack of uniformity; therefore, for future development, we suggest that individual registries should also record the actual estimated height of a fall in meters (as a continuous variable), so that an analysis can be performed.

We are aware that the best way of assessing the GOS as a measure of disability (morbidity), is after at least six months, but since this was not considered feasible, registration at discharge from the hospital was chosen as the endpoint. The assessment of GOS as an outcome measure at discharge from a trauma centre in trauma systems based on early discharge to specialised rehabilitation services represent a limitation, and the GOS endpoint should therefore be interpreted with caution. Nevertheless, assessment of this endpoint data can be used as a rough estimate of the amount of care needed for a patient beyond the acute hospital stay. Although GOS was developed for patients with head injuries, it represents a rough disability outcome score, and as such it should be possible to use for assessing all trauma patients.

The revised Utstein Template considers time to normalised arterial base excess as a measurement of importance for evaluating the total quality of trauma treatment. This data is not a precisely timed measurement and will probably vary from hospital to hospital. However, collecting it is considered valuable and might be important for future comparisons. The inclusion of this data also has some educational reasons; the trauma centres should be encouraged to include this value regularly in the trauma patients’ charts.

The expert panel recommends reporting the first EI, and the time to first EI, disregarding the fact that some patients may have received several of the listed interventions. The first intervention is, by clinical judgement, the most important one for comparison. This does not preclude trauma registries from documenting each of these interventions separately, or even each operation that is performed. However, the type of first EI and the time elapsed prior to its application is of significant importance; for this reason, this measure was chosen as part of the core dataset. For the purpose of comparison, focusing on the first EI performed in the current system will be adequate. However, if comparisons based on all EI are desired, the present recommendation will not be adequate. Comparison of systems based on all interventions can only be used if it is coupled with data on the time of each intervention. This will be difficult to use if the order in which the interventions occur is not known. If the first intervention is not among those listed under the EIs in the template, time to intervention is not important in our context (i.e., there will be enough time for thorough surgical preparation), and it should therefore not be part of the Utstein core dataset.

The panel acknowledges that no indicator has 100% validity and careful judgement is therefore needed. For example, in some selected cases earlier emergency surgery may imply both later CT imaging and better practice. Definition of data variables is a complex and ongoing process and in order to widen the implementation of the core dataset, facilitate participation in European trauma audit and comparisons and increase the quality of the next updates of the core dataset, we encourage all readers to ask for clarifications and point out potential improvements.

One concern with the Utstein Template is implementation, since none of the participants from the previous Utstein trauma process changed their registries to accommodate the recommendations of that meeting. The solution for the revised template was to develop a letter of consent that all members of the expert panel signed where they agreed on implementing the core data in the revised Utstein Template [see Additional file 1]. To further facilitate implementation, a complete user manual with core data variables and variable definitions will be available free of charge, at: http://www.scantem.org, https://www.tarn.ac.uk, http://www.traumaregister.de/de/index.htm, http://www.pprg.infoteca.it/ritg.

Conclusion
This paper represents a major step in perfecting the Utstein Template for Uniform Reporting of Data following Major Trauma, making the core data variables more uniform and applicable. Collecting this core dataset should be a basic component of all future studies on trauma care, and a uniform dataset such as this, will facilitate accurate description of the patient population and allow comparisons of outcome from trauma systems. It is extremely important that the data variables are collected.
in a uniform manner. For this reason, each variable and response category has been specifically defined in a way that is designed to promote the collection and reporting of a comparable core dataset. A letter of consent has been signed by the expert panel, where the participants of this consensus process agreed to implement the inclusion and exclusion criteria and core data variables in their respective systems and registries.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
KGR and HML designed the study and organised the consensus meetings. TJC was the chairman of the consensus process, while LH and OR were co-chairmen. The core dataset was developed by the expert panel. KGR, TJC, RL, SDB, OR, LH, PAS and HML wrote the article. The rest of the panelists read the manuscript once, made comments to it and approved the manuscript. All authors read and approved the final version of the manuscript.

Additional material

Addtional file 1
Letter of Consent. The letter provided was signed by the Utstein TCD expert panel. With this letter, the expert panel members confirm that they will implement the core data agreed upon.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1757-7241-16-7-S1.pdf](http://www.biomedcentral.com/content/supplementary/1757-7241-16-7-S1.pdf)

Acknowledgements
TCD = "Trauma Core Data". The members of the Utstein TCD expert panel were as follows:

Stefano Di Bartolomeo (Italy); Maaret Castrén (Sweden); Erika F. Christiansen (UK); Antoinette Edwards (UK); Torsten Eken (Norway); Emestina Gomes (Portugal); Lauri Handolin (Finland); Morten Hestnes (Norway); Lena Klarin (Sweden); Morten Schultz Larsen (Denmark); Rolf Lefering (Germany); Ari Leppäniemi (Finland); Hans Morten Lossius (Norway); Torben Wismorg (Norway); Maralyn Woodford (UK).

We acknowledge the Working Party of the ITACCS for developing the first Utstein Template, as well as the previous work of the EuroTARN group. We also wish to thank Jens Lauritsen MD, PhD, Accident Analysis Group, Odense University Hospital, Odense, Denmark and Professor Eldar Sareide, Departments of Emergency and Intensive Care Medicine, Division of Acute Care Medicine, Stavanger University Hospital, Stavanger, Norway for their contributions and support to this project.

Funding: The consensus process was funded by the Norwegian Air Ambulance Foundation, Drøbak, Norway (KGR, PhD scholarship, and expenses for expert panel meetings), and the Laerdal Foundation for Acute Medicine, Stavanger, Norway (expenses for expert panel meetings). The sponsors did not influence the final decision to submit for publication, nor did they have any role in the writing process, other than facilitating the meetings.

References
18. Skaga NO, Eken T, Steen PA: Assessing quality of care in a trauma referral center: benchmarking performance by...
24. Trauma Audit and Research Network: TARN. [https://www.tarn.ac.uk/]
27. SCANTEM: Scandinavian Networking Group for Trauma and Emergency Management. [http://www.scantem.org/]
32. Italian National Registry of Major Injuries: Registro Intraospedaliere Regionale Traumi Gravi. [http://www.prpp.infotecnica.it/ritg/]
36. American College of Surgeons: National Trauma Data Bank; National Trauma Data Standard, Data Dictionary v. 1.2. [http://www.nstdsdictionary.org/]


Collecting core data in severely injured patients using a consensus trauma template: an international multicentre study

Kjetil Gorseth Ringdal1,2,3*, Hans Morten Lossius1,4, J Mary Jones1,5, Jens M Lauritsen6,7, Timothy J Coats8,9, Cameron S Palmer10, Rolf Lefering11,12, Stefano Di Bartolomeo13,14, David J Dries15,16 and Kjetil Søreide4,17, for The Utstein Trauma Data Collaborators

Abstract

Introduction: No worldwide, standardised definitions exist for documenting, reporting and comparing data from severely injured trauma patients. This study evaluated the feasibility of collecting the data variables of the international consensus-derived Utstein Trauma Template.

Methods: Trauma centres from three different continents were invited to submit Utstein Trauma Template core data during a defined period, for up to 50 consecutive trauma patients. Directly admitted patients with a New Injury Severity Score (NISS) equal to or above 16 were included. Main outcome variables were data completeness, data differences and data collection difficulty.

Results: Centres from Europe (n = 20), North America (n = 3) and Australia (n = 1) submitted data on 965 patients, of whom 783 were included. Median age was 41 years (interquartile range (IQR) 24 to 60), and 73.1% were male. Median NISS was 27 (IQR 20 to 38), and blunt trauma predominated (91.1%). Of the 36 Utstein variables, 13 (36%) were collected by all participating centres. Eleven (46%) centres applied definitions of the survival outcome variable that were different from those of the template. Seventeen (71%) centres used the recommended version of the Abbreviated Injury Scale (AIS). Three variables (age, gender and AIS) were documented in all patients. Completeness > 80% was achieved for 28 variables, and 20 variables were > 90% complete.

Conclusions: The Utstein Template was feasible across international trauma centres for the majority of its data variables, with the exception of certain physiological and time variables. Major differences were found in the definition of survival and in AIS coding. The current results give a clear indication of the attainability of information and may serve as a stepping-stone towards creation of a European trauma registry.

Introduction

Major trauma is a leading cause of death and disability around the world [1], and it accounts for approximately 10% of the world’s deaths. Globally, unintentional injuries are ranked as the sixth leading cause of death and the fifth leading cause of moderate and severe disability [2]. The introduction of regionalised trauma systems has the potential to reduce preventable deaths [3], but an improved understanding of the benefits and limitations of different trauma care systems requires comparison across systems [4]. However, it has been shown that the datasets of existing trauma registries frequently lack compatible definitions of common data variables [5-9]. Consequently, the comparison and interpretation of trauma system outcomes has been hampered [10]. The lack of dataset uniformity poses substantial challenges to initiatives seeking to assess the quality of healthcare systems [11]. Several regions, particularly in North America, have implemented systematic documentation of trauma care and trauma system performance [12]. However, such documentation is limited in Europe [5,13,14], where no joint trauma registry exists [5,15].

* Correspondence: kjetil.ringdal@norskluftambulanse.no
1Department of Research, Norwegian Air Ambulance Foundation, Holtverveien 24, N-1440 Drøbak, Norway
Full list of author information is available at the end of the article.

http://ccforum.com/content/15/5/R237

© 2011 Ringdal et al; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
A recent European collaboration (the EuroTARN Group) assessed the potential of creating a data collection trial among a number of trauma registries in Europe, and the potential for comparing summary data and crude mortality rates [5]. Due to differences between trauma registries, the collaboration recognised that meaningful outcome comparisons were not possible. Similar conclusions were also reached in a contemporary Scandinavian report [9].

To address the discrepancies raised in these reports, members of the German, Italian, Scandinavian and UK trauma registries [15,16] performed an expert panel consensus process to develop a core set of uniform patient-level data for documenting and reporting trauma incidents. The resulting template, the Utstein Trauma Template [15-17], consists of recommended eligibility criteria and a set of 36 core data variables with four subsidiary variables.

The aim of the current study was to evaluate the feasibility of collecting patient-level data for severely injured patients across trauma centres using the Utstein Trauma Template variables as a standard.

Materials and methods

Study design

The study was a prospective international multicentre feasibility study, in which each participating institution was asked to collect and code up to 50 consecutively hospitalised trauma patients during the study period. The reporting of this study aims at conforming to the STROBE statement for reporting observational studies [18].

Participants

The primary focus was on inviting trauma registries from a mix of small, medium and large volume European trauma centres. However, to ensure that a degree of valid worldwide comparability was assessed, centres from North America and Australia were also invited.

Trauma centres were invited using a standardised open letter sent by email. For centres that agreed to participate, three reminder emails were sent to those that had not submitted data within the deadline. No follow-up was performed for the institutions that did not respond to the first invitation letter.

Patients

Trauma centres were asked to include directly and consecutively admitted trauma patients with a New Injury Severity Score (NISS) [19] ≥16 who presented between 1 September 2009 and 30 November 2009. Patients were excluded if they were transferred to the hospital, admitted to the hospital > 24 hours after injury, or if they were declared dead before hospital arrival or with no signs of life upon hospital arrival and no response to initial hospital resuscitation. Patients with asphyxia or drowning injuries and patients who had burns as the predominant injury were also excluded [16].

Data variables

Participants were asked to collect all the data variables of the Utstein Trauma Template [16] and to fill out and return a self-administered questionnaire (Additional file 1). Using the questionnaire, the centres were asked to report the data variables that they were able to collect, whether their data variable definitions deviated from the definitions of the template, and if they experienced any data collection difficulties. Additional comments could be made for each variable. The centres were asked to grade all injuries according to the Abbreviated Injury Scale (AIS) 2005 or 2005-update 2008 [20], reporting the whole seven digit AIS code.

The main outcome measures were the data completeness of information at the patient and variable levels, the discrepancy for data variable definitions, and the difficulty of data collection. The completeness of the patient-level data was measured on the basis of reported values, while unknown and blank values were considered as missing values. The discrepancy for definitions and the collection difficulty were assessed from the associated questionnaire and were based on “yes”, “no” and “unknown” answers.

Data collection

Patient-level data were collected using the local hospital-based or regional trauma registries. For participants without a suitable registry, an electronic database was provided by the investigators. Centres that did not return the questionnaires, provide the patient-level data, or return a consent form were excluded from the study.

Study size

The sample size of 50 cases per centre was chosen as a pragmatic size to balance the workload imposed by the study while providing a reasonable number of patients for computing completeness proportions. Low-volume trauma centres that were not able to collect 50 cases within the timeframe were asked to submit three months of hospitalised patients.

Ethics

The study was approved by the Regional Committee for Medical and Health Research Ethics of Southeast Norway and the Norwegian Social Sciences Data Services. The exported datasets were required to not contain any direct or indirect patient identifiable data. Dates and times were not permitted in the submitted datasets, and patients could not be marked with a reference number.
that could be linked to a patient number from the submitting centre. The centres were required to return a written and signed consent form stating that the participation and data sharing was in compliance with their own specific institutional and/or national legal frameworks and data protection requirements.

Statistical methods
Continuous data that were not normally distributed are presented using median and interquartile range (IQR) and analysed using non-parametric techniques. Categorical data are presented as counts and proportions. The completeness of the Utstein core data variables are presented as counts and proportions with 95% confidence intervals (using Wilson’s method [21]), and the completeness of each dataset was judged by the number of centres with complete patient data using percentile levels (50th, 75th and 100th). The desired goal of data completeness was set at ≥80%.

Data were analysed using IBM SPSS version 18 (IBM Company, Chicago, IL, USA) and Stata/SE version 11.1 (StataCorp LP, College Station, TX, USA).

Results
Participating centres
Of the 42 centres invited, 10 never responded, four centres declined to participate but never submitted the requested material. Twenty-four of the invited centres (57%) participated in the study (Figure 1), of which nine had been part of the Utstein Template development process.

In total, 14 nations were represented. Participation amongst Scandinavian invitees was 64%, participation amongst European invitees outside Scandinavia was 61%, and participation amongst North American/Australian centres was 40%.

Two participants were large, multi-institutional trauma registries that represented collaborations of hospitals (152 and 120 hospitals, respectively), and 20 participants were individual hospitals with a hospital-based registry (Table 1). Two participants did not have a registry prior to the initiation of the study. The two multinational trauma registries and eight centres with a hospital-based registry had fully or partially implemented the Utstein variables prior to the initiation of the study. The remaining 14 centres only collected the Utstein variables in the trauma registries.

Patient characteristics
In total, data from 965 patients were submitted. Of these, 182 (19%) were excluded for not meeting the study selection criteria (Figure 1). AIS codes were missing for 12 patients, 94 patients had a NISS < 16, and 76 patients were transferred to the reporting hospital. Therefore, 783 (81%) patients were available for analyses, with 623 (80%) patients from European centres.

Patient characteristics are summarised in Table 2. The majority of the patients were male (73.1%), and the median age was 41.0 years (IQR 24 to 60). Blunt trauma predominated (91.1%), while traffic accidents (53.1%) and high-energy falls (19.3%) were the most prominent injury mechanisms. The median NISS was 27.0 (IQR 20 to 38), and the reported death rate was 14.0%.

Data variables collected by centres
Of the 36 Utstein variables, 13 (36%) variables were collected in all 24 centres (Table 3). The variable that was recorded by the fewest centres was “Time Until Normal Arterial Base Excess”, which was recorded by 17 participants (70.8%) with a completion level of 48.2% (Figure 2 and Additional file 2). Of all the Utstein variables, four (11%) variables did not deviate from the template’s definitions in any of the centres (Table 3). Several centres had variable definitions that differed from the definitions of the Utstein Template. The most heterogeneously defined variable was “Survival Status” (the Utstein recommendation is outcome at Day 30 after injury [11]), and 11 (46%) centres used different definitions (Table 3): six used outcome at end of acute care stay, three used the in-hospital 30-day outcome, and two used the outcome at the end of total somatic stay (including rehabilitation). All centres used the AIS system for anatomical severity scoring. However, only 17 (71%) of the centres used the versions recommended. Two centres submitted the single-digit AIS severity codes, excluding the six-digit injury descriptor.

Only two variables, “Gender” and “Age”, were collected from all centres without difficulty. The variable that was most frequently reported to be difficult to collect was “Pre-Hospital Respiratory Rate”, which eight centres (35%) reported as difficult (Table 3).

Completeness of patient-level core data
The levels of completion for each of the Utstein variables are shown in Figure 2 and Additional file 2. Some centres declined to record specific variables but nevertheless submitted data on those variables for some patients. After exclusion of these datasets, the results showed that 20 Utstein core variables were > 90% complete. Of these, three variables (age, gender, and AIS) were 100% complete. Twenty-eight data variables were >
80% complete. Eight variables had completeness levels that were below the desired 80% threshold (Figure 2). The variables “Time Until Normal Arterial Base Excess”, “Arterial Base Excess”, and “Pre-Hospital Respiratory Rate”, had the lowest levels of completeness.

For reporting pre- and in-hospital SBP and RR values, the Utstein Template recommends the use of clinical categories (based on the Revised Trauma Score (RTS) categories [22]) when continuous values are missing [15]. This is illustrated in the results presented in Figure 2 and Additional file 2. When the continuous and categorical values of pre-hospital SBP and RR were combined, the completeness increased by 8.9% and 23.2%, respectively (Figure 3). The equivalent in-
hospital completeness levels showed an increase of 1.9% and 17.6%.

Discussion

The current international multicentre study demonstrated acceptable feasibility and completeness in reporting trauma data using a common template. For the majority of variables, the data collection was sufficient, while some areas in need for improvement were identified. The feasibility of bearing this project to fruition may serve as a stepping-stone towards establishment of a common pan-European trauma registry. However, some results deserve further discussion.

This study demonstrated that the data for 28 (78%) of the Utstein variables were > 80% complete, and that the data for 20 (56%) variables were > 90% complete. The pre-hospital SBP and RR values were less complete than were the equivalent in-hospital values. This result is consistent with findings from Arbabi et al. [23], who found that pre-hospital and admission SBP values were recorded for 35% and 67% patients, respectively. In cases with missing continuous values, the Utstein Template recommends documenting the SBP and RR values as RTS categories [15,16]. This recommendation is not merely a mathematical consideration; it has a practical sense because clinical categories can be reasonably approximated by palpation of the patient’s pulses and by chest examination. In the present study, the combination of the continuous and categorical SBP and RR values resulted in increased completeness compared to the sole use of continuous values (Figure 3). Although categorising continuous data may result in loss of precision and power in addition to other methodological challenges [24,25], the use of the clinical categories provides an undeniable advantage over not having data.

Table 1 Characteristics of participating centres (n = 24)

<table>
<thead>
<tr>
<th>Centre characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection for the project, n (%)</td>
<td></td>
</tr>
<tr>
<td>In own registry</td>
<td>19 (79.2%)</td>
</tr>
<tr>
<td>In database designed for purpose</td>
<td>5 (20.8%)</td>
</tr>
<tr>
<td>Utstein data variables implemented prior to study initiation, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>10 (41.7%)</td>
</tr>
<tr>
<td>No</td>
<td>14 (58.3%)</td>
</tr>
<tr>
<td>Type of pre-study registry, n (%)</td>
<td></td>
</tr>
<tr>
<td>Local/regional</td>
<td>20 (83.3%)</td>
</tr>
<tr>
<td>National/multi-national</td>
<td>2 (8.3%)</td>
</tr>
<tr>
<td>No registry</td>
<td>2 (8.3%)</td>
</tr>
<tr>
<td>Number of patients with ISS &gt; 15 received in 2008, median (IQR)</td>
<td>191 (110 to 490); n = 17</td>
</tr>
<tr>
<td>Number of patients with NISS &gt; 15 in 2008, median (IQR)</td>
<td>200 (78 to 1794); n = 10</td>
</tr>
<tr>
<td>Number of hospitals the centre/registry represents, median (range)</td>
<td>1 (1 to 152)</td>
</tr>
</tbody>
</table>

IQR, Interquartile range; ISS, Injury Severity Score; NISS, New Injury Severity Score.

All centres reported injuries according to the AIS system, although injury documentation standards varied. Even though the majority of participating institutions used the AIS dictionaries recommended, nearly 30% did not. Several recent studies have identified differences between the AIS 1998 and 2005/2008 dictionaries in terms of the number of patients classified as ‘major trauma’ [26-28], illustrating that injury data collected using different AIS dictionaries cannot be directly compared. When comparing outcomes, Injury Severity Score (ISS) [29] or NISS values, AIS dictionary differences could affect the discrimination between severely and less severely injured patients across national and international registries. In light of the recent literature, it is not clear whether parallel coding using the AIS 1998 and AIS 2005/2008 versions should be recommended in order to enable comparisons. However, a solution to overcome the limitation of the existing mapping tool in the AIS dictionary [30] may be a newly developed AIS98 to AIS08 mapping tool [30].

The Utstein Template recommends the use of the short-term outcome variable ‘30-day survival’, which is a mortality indicator that is also applied in other fields such as stroke and acute myocardial infarction [31]. The definitions of the survival outcome variable differed across the participating centres included in the current study. Some centres evaluated short-term outcome based on hospital administrative data, which resulted in the use of in-hospital survival or in-hospital 30-day survival. Others used 30-day survival regardless of whether the patient was still hospitalised. These differences may result in unfavourable biases when trauma care is compared. The use of in-hospital 30-day survival can be particularly problematic with short length of hospital stay or increased tendencies for transfer of patients between facilities. Thus, a greater proportion of deaths within 30
days of injury may be missed if only ‘in-hospital deaths’ are considered [11,15,31-33]. The endpoints ‘in-hospital survival’ and ‘30-day survival’ should both be considered included in the Utstein Template until the health care systems have matured to the point where data about 30-day survival status are easily obtainable.

This study does have associated limitations. First, the process applied for identifying centres for invitation was subjective and not standardised. Participating centres may be more likely to comply with the Utstein Template or better able to collect and report data requested. Second, the 10 centres not responding to invitations and the 4 that agreed to participate but never submitted the requested materials were not further contacted. Thus, we cannot preclude the possibility that they found collection of the dataset too difficult or time consuming.

Third, some institutions had already integrated the Trauma Template variables in their trauma documentation protocols and registries prior to the start of the current study, while others only collected these data for the study. Implementation of the template across all centres should have yielded a higher degree of completeness for some data sets. Fourth, participation from North America and Australia was low. However, because of the formalised criteria with which trauma care in American centres is reviewed, there is a greater homogeneity among these centres and data collection. Thus, despite the small number of hospitals, inclusion of three leading centres from the United States gives a good sampling of North American practice. Fifth, the desired goal of completeness (> 80%) used in this study is an arbitrarily chosen threshold. No justifications or guidelines for the acceptability of missing data in registry studies (for example, prognostic studies) exist [personal communication with Professor Douglas G. Altman, University of Oxford, UK]. Thus, the threshold value was a choice based on consensus among the authors. Finally, the template allows some data fields to be left blank when a data variable is unknown or not documented. Leaving a data field blank can make it more difficult to estimate the exact completeness or perform comparative analyses (that is, the exact cause of leaving a data field blank could be “not measured”, “forgotten” or “unknown”).

This study demonstrates that considerable support exists for the development of an international uniform mandatory core dataset that can be the basis of a European trauma registry. However, several steps still remain. The current Utstein Trauma Template variables and definitions could be further improved before collaborative research on the comparison of trauma care performance is initiated on a larger scale. Hopefully, the results from this study will contribute to improvements. Indeed, at the time of the development of the Utstein

Table 2 Characteristics of the included trauma patients (n = 783)

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR)</td>
<td>41 (24 to 60)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>572 (73.1%)</td>
</tr>
<tr>
<td>Female</td>
<td>211 (26.9%)</td>
</tr>
<tr>
<td>Dominating type of injury, n (%)</td>
<td></td>
</tr>
<tr>
<td>Blunt</td>
<td>713 (91.1%)</td>
</tr>
<tr>
<td>Penetrating</td>
<td>68 (8.7%)</td>
</tr>
<tr>
<td>Missing data</td>
<td>2 (0.2%)</td>
</tr>
<tr>
<td>Mechanism of injury, n (%)</td>
<td></td>
</tr>
<tr>
<td>Traffic: motor vehicle accident (excluding motorcycle)</td>
<td>154 (19.7%)</td>
</tr>
<tr>
<td>Traffic: motorcycle accident</td>
<td>114 (14.6%)</td>
</tr>
<tr>
<td>Traffic: bicycle accident</td>
<td>48 (6.1%)</td>
</tr>
<tr>
<td>Traffic: pedestrian</td>
<td>68 (8.7%)</td>
</tr>
<tr>
<td>Traffic: other</td>
<td>31 (4.0%)</td>
</tr>
<tr>
<td>Shot</td>
<td>36 (4.6%)</td>
</tr>
<tr>
<td>Stabbed</td>
<td>31 (4.0%)</td>
</tr>
<tr>
<td>Struck or hit by blunt object</td>
<td>33 (4.2%)</td>
</tr>
<tr>
<td>Low energy fall</td>
<td>87 (11.1%)</td>
</tr>
<tr>
<td>High energy fall</td>
<td>151 (19.3%)</td>
</tr>
<tr>
<td>Other</td>
<td>22 (2.8%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>5 (0.6%)</td>
</tr>
<tr>
<td>Missing data</td>
<td>3 (0.4%)</td>
</tr>
<tr>
<td>Injuries grouped by AIS body regions, n (%)</td>
<td></td>
</tr>
<tr>
<td>Head</td>
<td>1,148 (26.4%)</td>
</tr>
<tr>
<td>Face</td>
<td>407 (9.4%)</td>
</tr>
<tr>
<td>Neck</td>
<td>16 (0.4%)</td>
</tr>
<tr>
<td>Thorax</td>
<td>713 (16.4%)</td>
</tr>
<tr>
<td>Abdomen</td>
<td>252 (5.8%)</td>
</tr>
<tr>
<td>Spine</td>
<td>426 (9.8%)</td>
</tr>
<tr>
<td>Upper extremity</td>
<td>388 (8.9%)</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>532 (12.3%)</td>
</tr>
<tr>
<td>External and other</td>
<td>85 (2.0%)</td>
</tr>
<tr>
<td>Missing data</td>
<td>375 (8.6%)</td>
</tr>
<tr>
<td>Injuries grouped by AIS severity levels, n (%)</td>
<td></td>
</tr>
<tr>
<td>AIS 1 to 3</td>
<td>3,554 (81.9%)</td>
</tr>
<tr>
<td>AIS 4 to 6</td>
<td>772 (17.8%)</td>
</tr>
<tr>
<td>AIS 9 (unknown)</td>
<td>16 (0.3%)</td>
</tr>
<tr>
<td>NISS groups, n (%)</td>
<td></td>
</tr>
<tr>
<td>16 to 24</td>
<td>313 (40.0%)</td>
</tr>
<tr>
<td>25 to 40</td>
<td>283 (36.1%)</td>
</tr>
<tr>
<td>41 to 56</td>
<td>123 (15.7%)</td>
</tr>
<tr>
<td>57 to 75</td>
<td>64 (8.2%)</td>
</tr>
<tr>
<td>Survival status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Died</td>
<td>110 (14.0%)</td>
</tr>
<tr>
<td>Survived</td>
<td>621 (79.3%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>43 (5.5%)</td>
</tr>
<tr>
<td>Missing data</td>
<td>9 (1.1%)</td>
</tr>
</tbody>
</table>

AIS, Abbreviated Injury Scale; IQR, Interquartile range; NISS, New Injury Severity Score
Table 3 Number and proportion of collected Utstein variables, differences in variable definitions, and data collection difficulties

<table>
<thead>
<tr>
<th>Core data variable</th>
<th>Centres collecting this data variable n (%)</th>
<th>Applied a different definition n (%)</th>
<th>Data variable was difficult to collect n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>24 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>in-hospital SBP</td>
<td>24 (100%)</td>
<td>0</td>
<td>1 (4.2%)</td>
</tr>
<tr>
<td>Hospital length of stay</td>
<td>24 (100%)</td>
<td>0</td>
<td>1 (4.2%)</td>
</tr>
<tr>
<td>Age</td>
<td>24 (100%)</td>
<td>1 (4.2%)</td>
<td>0</td>
</tr>
<tr>
<td>Dominating type</td>
<td>24 (100%)</td>
<td>2 (8.3%)</td>
<td>1 (4.2%)</td>
</tr>
<tr>
<td>Intention of injury</td>
<td>24 (100%)</td>
<td>2 (8.3%)</td>
<td>3 (12.5%)</td>
</tr>
<tr>
<td>Discharge destination</td>
<td>24 (100%)</td>
<td>4 (16.7%)</td>
<td>2 (8.3%)</td>
</tr>
<tr>
<td>Mechanism of injury</td>
<td>24 (100%)</td>
<td>7 (29.2%)</td>
<td>3 (12.5%)</td>
</tr>
<tr>
<td>Abbreviated Injury Scale</td>
<td>24 (100%)</td>
<td>8 (33.3%)</td>
<td>1 (4.2%)</td>
</tr>
<tr>
<td>Survival status</td>
<td>24 (100%)</td>
<td>11 (45.8%)</td>
<td>4 (16.7%)</td>
</tr>
<tr>
<td>in-hospital RR</td>
<td>23 (95.8%)</td>
<td>0</td>
<td>6 (25.0%)</td>
</tr>
<tr>
<td>Pre-hospital SBP</td>
<td>23 (95.8%)</td>
<td>1 (4.3%)</td>
<td>5 (21.7%)</td>
</tr>
<tr>
<td>Pre-hospital GCS</td>
<td>23 (95.8%)</td>
<td>2 (8.7%)</td>
<td>5 (21.7%)</td>
</tr>
<tr>
<td>Pre-hospital RR</td>
<td>23 (95.8%)</td>
<td>2 (8.7%)</td>
<td>8 (34.8%)</td>
</tr>
<tr>
<td>Pre-hospital GCS motor component</td>
<td>22 (91.7%)</td>
<td>2 (9.1%)</td>
<td>6 (27.3%)</td>
</tr>
<tr>
<td>in-hospital GCS</td>
<td>22 (91.7%)</td>
<td>2 (9.1%)</td>
<td>1 (4.5%)</td>
</tr>
<tr>
<td>Pre-hospital cardiac arrest</td>
<td>22 (91.7%)</td>
<td>3 (13.6%)</td>
<td>1 (4.5%)</td>
</tr>
<tr>
<td>Pre-injury ASA-PS classification</td>
<td>22 (91.7%)</td>
<td>7 (31.8%)</td>
<td>6 (27.3%)</td>
</tr>
<tr>
<td>in-hospital GCS motor component</td>
<td>21 (87.5%)</td>
<td>2 (9.5%)</td>
<td>1 (4.8%)</td>
</tr>
<tr>
<td>Days on ventilator</td>
<td>21 (87.5%)</td>
<td>5 (23.8%)</td>
<td>5 (23.8%)</td>
</tr>
<tr>
<td>INR</td>
<td>20 (83.3%)</td>
<td>3 (15.0%)</td>
<td>1 (5.0%)</td>
</tr>
<tr>
<td>GOS score at discharge</td>
<td>20 (83.3%)</td>
<td>3 (15.0%)</td>
<td>5 (25.0%)</td>
</tr>
<tr>
<td>Pre-hospital SBP-clinical category *</td>
<td>19 (79.2%)*</td>
<td>1 (5.3%)*</td>
<td>1 (5.3%)*</td>
</tr>
<tr>
<td>Arterial base excess</td>
<td>19 (79.2%)</td>
<td>1 (5.3%)</td>
<td>3 (15.8%)</td>
</tr>
<tr>
<td>In-hospital RR-clinical category *</td>
<td>18 (75.0%)*</td>
<td>2 (11.1%)*</td>
<td>2 (11.1%)*</td>
</tr>
<tr>
<td>Pre-hospital RR-clinical category *</td>
<td>18 (75.0%)*</td>
<td>2 (11.1%)*</td>
<td>4 (22.2%)*</td>
</tr>
<tr>
<td>In-hospital SBP-clinical category *</td>
<td>17 (70.8%)*</td>
<td>1 (5.9%)*</td>
<td>1 (5.9%)*</td>
</tr>
<tr>
<td>System characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-hospital transfer</td>
<td>24 (100%)</td>
<td>1 (4.2%)</td>
<td>2 (8.3%)</td>
</tr>
<tr>
<td>Transportation type</td>
<td>24 (100%)</td>
<td>4 (16.7%)</td>
<td>2 (8.3%)</td>
</tr>
<tr>
<td>Type of first key emergency intervention</td>
<td>24 (100%)</td>
<td>4 (16.7%)</td>
<td>2 (8.3%)</td>
</tr>
<tr>
<td>Highest level of in-hospital care</td>
<td>23 (95.8%)</td>
<td>1 (4.3%)</td>
<td>4 (17.4%)</td>
</tr>
<tr>
<td>Pre-hospital airway management</td>
<td>23 (95.8%)</td>
<td>3 (13.0%)</td>
<td>4 (17.4%)</td>
</tr>
<tr>
<td>Trauma team activation</td>
<td>22 (91.7%)</td>
<td>2 (9.1%)</td>
<td>3 (13.6%)</td>
</tr>
<tr>
<td>Time from alarm until hospital arrival</td>
<td>22 (91.7%)</td>
<td>2 (9.1%)</td>
<td>5 (22.7%)</td>
</tr>
<tr>
<td>Highest level of pre-hospital care provided</td>
<td>22 (91.7%)</td>
<td>4 (18.2%)</td>
<td>4 (18.2%)</td>
</tr>
<tr>
<td>Type of pre-hospital airway management</td>
<td>20 (83.3%)</td>
<td>4 (20.0%)</td>
<td>4 (20.0%)</td>
</tr>
<tr>
<td>Process mapping data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time until first CT scan</td>
<td>23 (95.8%)</td>
<td>4 (17.4%)</td>
<td>7 (30.4%)</td>
</tr>
<tr>
<td>Time from alarm until arrival at scene</td>
<td>22 (91.7%)</td>
<td>1 (4.5%)</td>
<td>6 (27.3%)</td>
</tr>
<tr>
<td>Time until first key emergency intervention</td>
<td>22 (91.7%)</td>
<td>4 (18.2%)</td>
<td>4 (18.2%)</td>
</tr>
<tr>
<td>Time until normal arterial base excess</td>
<td>17 (70.8%)</td>
<td>2 (11.8%)</td>
<td>6 (35.3%)</td>
</tr>
</tbody>
</table>

* The clinical categories relate to cases for which continuous data were not submitted.

ASA-PS, American Society of Anesthesiologists Physical Status; EMS, emergency medical service; GCS, Glasgow Coma Scale; GOS, Glasgow Outcome Scale; LOS, length of stay; RR, respiratory rate; SBP, systolic blood pressure
Figure 2 Completeness of the Utstein core variables among the participating centres. Current completeness of Utstein core variables (n = 24 centres). The proportion of centres collecting each variable, and the proportion of eligible patients with reported information (with 95% CI) are shown.
Template, the balance between ‘desirability’ and ‘collectability’ of a variable was probably in favour of the former because there were no objective data on ‘collectability’. This study has identified variables that are particularly difficult to collect. In particular, the collection of “Arterial Base Excess”, and process data like “Time Until Normal Arterial Base Excess” needs to be reconsidered or even excluded from the template, while uniform survival outcome variables and type of AIS coding systems used, should be further agreed upon. Additional studies should review the propriety of some of the variables. Furthermore, the data variables should be evaluated with regard to inter-rater reliability [34].

The template was primarily developed for patients who were directly admitted to a trauma centre. A more complete assessment of the performance of the entire trauma system [35] will need to include transferred patients. Exclusion of transferred patients may strongly influence the results when hospitals with large proportions of transferred patients are included.

In order to further develop an international core data set, a consensus-driven revision of the Utstein Trauma Template, with representatives from multiple continents, should be initiated. The results of the current study will be valuable for such a revision.

**Conclusions**

The study shows that 78% of the data variables of the Utstein Trauma Template were > 80% complete. Difficulty with collecting time variables and a lack of uniformity in the use of outcome variables and injury scoring systems across international trauma institutions were found. Overall, the feasibility of collecting most of the core data was demonstrated across several registries and countries.

**Key messages**

- The use of the common trauma template was feasible across international registries for the majority of the data variables.
- A lack of uniformity in the use of outcome variables and injury scoring systems across international trauma institutions mandate a need for better standardisation.
- The current results may serve as a stepping-stone towards creation of a European trauma registry.
Additional material

Additional file 1: The self-administered questionnaire. This file includes the self-administered questionnaire that was distributed to the participating centres.

Additional file 2: Completeness of the Utstein core variables among the participating centres. The file includes the number of centres collecting a data variable, completeness of patient data from the recording centres, and the number of centres with complete patient data by percentiles.

Abbreviations
AIS: Abbreviated Injury Scale; QRR: interquartile range; ISS: Injury Severity Score; NISS: New Injury Severity Score; RR: respiratory rate; RTS: Revised Trauma Score; SBP: systolic blood pressure.

Acknowledgements
The Utstein Trauma Data Collaborators
The following people are acknowledged for their assistance with the data collection:

Australia: The Royal Children's Hospital Melbourne, Victoria: Cameron S. Palmer.

Denmark: Copenhagen University Hospital, Rigshospitalet, Copenhagen: Claus Falck Larsen; Odense University Hospital, Odense: Anne Ellegaard Hansen; Morten Schultz Larsen; Jens M. Lautrisen.

Finland: Meilahti Hospital, Helsinki University Hospital, Helsinki: Pia Hiennon, Ari Leppänen; Töölö Hospital, Helsinki University Hospital, Helsinki: Lauri Handolin.

Germany: Trauma Registry of the German Society of Trauma Surgery, Cologne: Rolf Lefering.

Greece: Asklepion Venal General Hospital, Athens: George Giannopoulos.


Lithuania: Vilnius University Children's Hospital, Vilnius: Ruta Kvederienė, Virginija Zilinskaite.

The Netherlands: Academic Medical Center, Amsterdam: J. Carel Goslings, Christiana Ponsen; VU University Medical Center, Amsterdam: Frank Bloemers, Geert Jan Giannakopoulos.

Norway: Oslo University Hospital-Ullevål, Oslo: Torstein Eken, Morten Hestnes, Olav Raisz, Nils Oddvar Skaga, Petter Andreas Steen; Stavanger University Hospital, Stavanger: Kjetil Søreide, Kjell Tjosevik; St. Olav University Hospital, Trondheim: Oddvar Ulleberg.

Portugal: Hospital de Santo António, Centro Hospitalar do Porto, Porto: Emestina Gomes.

Sweden: Karolinska University Hospital, Solna, Stockholm: Olaf Brattström; Södersjukhuset University Hospital, Stockholm: Maaret Castrèn, Rebecka Ringdal et al.

Switzerland: Centre Hospitalier Universitaire Vaudois, Lausanne: Catherine Hein.

United Kingdom: The Trauma Audit & Research Network, University of Manchester, Manchester: Timothy J. Coats, Antoinette Edwards, Thomas Lawrence, Fiona Lecky, Marlyn Woodford.

USA: Denver Health Medical Center, Denver, Colorado: Walter L. Biff, Craig Gravitz; Johns Hopkins Hospital Adult Trauma Center, Baltimore, Maryland: Adil H. Haeder, Kathy Noll; Regions Hospital, St. Paul, Minnesota and Gillette Children's Hospital, St. Paul, Minnesota: Heidi M. Altamirano, Sandy Brengman, David J. Dries, Michael D. McGonigal.

Supervisors
Professor Olav Raisz MD PhD, Nils Oddvar Skaga MD PhD and Professor Petter Andreas Steen MD PhD, Oslo University Hospital-Ullevål, Norway are acknowledged for their contributions to planning of the project and supervising KGR.

Discussions on completeness thresholds
We thank Professor Douglas G. Altman, Centre for Statistics in Medicine, University of Oxford, UK for valuable considerations on completeness thresholds.

AIS-to-NISS calculation system
Torsten Eken, MD PhD, Oslo University Hospital-Ullevål, Norway is acknowledged for providing a Microsoft Excel-based AIS-to-NISS conversion system.

Databases
The Italian Trauma Registry is acknowledged for providing the participating Italian hospitals with a trauma registry, and the UK Trauma Audit and Research Network is acknowledged for providing Copenhagen University Hospital with a registration tool. Acknowledgements are also extended to the EpiData Association, Odense, Denmark for providing a database to hospitals without an adequate registry.

Data storage
IT Department, Norwegian Air Ambulance Foundation, Drambak, Norway. PhD funding

The Norwegian Air Ambulance Foundation, Drambak, Norway funded the PhD scholarship of Kjetil G. Ringdal, MD.

Author details
1. Department of Research, Norwegian Air Ambulance Foundation, Holterveien 24, N-1440 Drambak, Norway. 2. Division of Emergencies and Critical Care, Oslo University Hospital-Ullevål, Kirkeveien 166, N-0450 Oslo, Norway. 3. Department of Radiology, Herlev Hospital, Herlev, Denmark. 4. Department of Anaesthesia and ICU, Azienda Ospedaliero-Universitaria di Udine, Piazzale Santa Maria della Misericordia, 33100 Udine, Italy. 5. Italian National Trauma Registry and Emilia-Romagna Trauma Registry, Department of Clinical Governance, Regional Health Agency, Viale Aldo Moro 21, 40127 Bologna, Italy. 6. Department of Surgery, Regions Hospital, 640 Jackson Street, St. Paul, MN 55101, USA. 7. Department of Surgery, University of Minnesota, 420 Delaware Street SE, Minneapolis, MN 55455, USA. 8. Department of Surgery, Stavanger University Hospital, Amauer Hansens vei 20, N-4011 Stavanger, Norway.

Authors’ contributions
KGR designed the study, prepared and analysed the data, drafted the results, and drafted the manuscript. HML designed the study, assisted in drafting the results, and drafted the manuscript. KS designed the study, assisted in drafting the results, and drafted the manuscript. JMJ designed the study, assisted in drafting the results, and drafted the manuscript. JML and CSP assisted in drafting the results, and assisted in drafting the manuscript. KGR designed the study, prepared and analysed the data, assisted in drafting the results, and assisted in drafting the manuscript. JML and CSP assisted in drafting the results, and assisted in drafting the manuscript. JML and CSP assisted in drafting the manuscript. All authors read and approved the final manuscript for publication. The Utstein Trauma Data Collaborators were invited to read and comment on the final manuscript draft.

Authors’ information
KGR, HML, JCC and RL are members of the European Trauma Registry Network’s working party for the establishment of a European trauma registry. KGR drafted the Utstein Trauma Template Data Dictionary and is a member of the working party for the establishment of the Norwegian National Trauma Registry. JCC is the chairman of the Trauma Audit and Research Network, UK. CSP chaired the working party of the National Trauma Registry Consortium in Australia and New Zealand for the formulation of a Binational Minimum Dataset (BMDS) and drafted the BMDS Data Dictionary in 2010. RL is co-chairman and statistician of the Trauma
Registry of the German Society of Trauma Surgery. SDP is the scientific director of the Italian National Trauma Registry.

Competing interests
The authors declare that they have no competing interests.

Received: 29 May 2011 Revised: 15 July 2011 Accepted: 12 October 2011

References
Abbreviated Injury Scale: not a reliable basis for summation of injury severity in trauma facilities

Kjetil G. Ringdal, RN, Cand. Mag., MD1,2,3,4,*
Nils Oddvar Skaga, MD, PhD5,6
Morten Hestnes, CRNA6
Petter Andreas Steen, MD, PhD2,3
Jo Røislien, M.Sc., PhD1,7
Marius Rehn, MD1,3,8
Olav Røise, MD, PhD2,3,4
Andreas J. Krüger, MD1,9,10
Hans Morten Lossius, MD, PhD1,11

1Department of Research, Norwegian Air Ambulance Foundation, Drøbak, Norway
2Division of Emergencies and Critical Care, Oslo University Hospital Ullevål, Oslo, Norway
3Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway
4Norwegian National Trauma Registry, Oslo University Hospital, Norway
5Department of Anaesthesiology, Division of Emergencies and Critical Care, Oslo University Hospital Ullevål, Oslo, Norway
6The Ullevål Trauma Registry, Department of Research and Development, Division of Emergencies and Critical Care, Oslo University Hospital Ullevål, Oslo, Norway
7Department of Biostatistics, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Norway
8Akershus University Hospital, Lørenskog, Norway
9Department of Anesthesia and Emergency Medicine, St. Olav’s University Hospital, Trondheim, Norway
10Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway
11Department of Surgical Sciences, Faculty of Medicine and Dentistry, University of Bergen, Bergen, Norway

*Correspondence to:
Dr. Kjetil G. Ringdal
Department of Research
Norwegian Air Ambulance Foundation
P.O. Box 94, N-1441 Drøbak, Norway
E-mail: kjetil.ringdal@norskluftambulanse.no
Tel: +47 64 90 44 44; Fax: +47 64 90 44 45
Abstract

Background: Injury severity is most frequently classified using the Abbreviated Injury Scale (AIS) as a basis for the Injury Severity Score (ISS) and New Injury Severity Score (NISS), which are used in outcome prediction. European trauma registries recommended the AIS 2008 edition, but the levels of inter-rater agreement and reliability of ISS and NISS, associated with its use, have not been reported.

Methods: Nineteen Norwegian AIS-certified trauma registry coders were invited to score 50 real, anonymized patient medical records using AIS 2008. Rater agreements for ISS and NISS were analysed using Bland-Altman plots with 95% limits of agreement (LoA). A clinically acceptable LoA range was set at ±9 units. The reliability was analysed using a two-way mixed model intraclass correlation coefficient (ICC) with a 95% confidence interval (CI) and hierarchical agglomerative clustering.

Results: Ten coders submitted their coding results. Of their AIS codes, 2189 (61.5%) agreed with a reference standard, 1187 (31.1%) real injuries were missed, and 392 non-existing injuries were recorded. All LoAs were wider than the predefined, clinically acceptable limit of ±9, for both ISS and NISS. The joint ICC (range) between each rater and the reference standard was 0.51 (0.29,0.86) for ISS and 0.51 (0.27,0.78) for NISS. The joint ICC (range) for inter-rater reliability was 0.49 (0.19,0.85) for ISS and 0.49 (0.16,0.82) for NISS. Univariate linear regression analyses indicated no significant relationship between the rater-against-reference ISS and NISS ICC values and the number of cases coded over the rater’s career or years of experience.

Conclusions: Based on AIS 2008, ISS and NISS were not reliable for summarizing anatomic injury severity in this study. This result limits their use as benchmarking tools for trauma system performance.
Introduction

Injury severity classification is considered a fundamental component of trauma outcome research and quality assessments. The Abbreviated Injury Scale (AIS)\(^1\), developed by the U.S. Association for the Advancement of Automotive Medicine (AAAM), is frequently used to classify severity of anatomic injuries. The AIS is the basis of several composite injury severity measurements, such as the Injury Severity Score (ISS)\(^2\) and New Injury Severity Score (NISS)\(^3\), and is one of the independent variables included in outcome prediction models such as the Trauma and Injury Severity Score (TRISS)\(^4,5\). Assigning AIS codes to patients with multiple injuries can be rater-subjective, because identical injuries can be given different codes\(^6\). Mackenzie et al. reported considerable variation in the mean number of AIS-scored injuries recorded per patient among raters using the AIS 1980 edition\(^7\). Physicians and nurses had higher intra-rater reliability than emergency medical technicians and nonclinical technicians. The inter-rater AIS score agreement was significantly higher for blunt than for penetrating injuries. Neale et al. found that although only 39\% of the AIS codes assigned by any two raters were identical, the inter-rater reliability for ISS was almost perfect\(^8\). In contrast, Zoltie et al. found a large variation in ISS, with a mere 28\% probability of agreement between two independent raters\(^9\).

A revised Utstein Trauma Template for uniform data reporting from patients subjected to major trauma was published recently\(^10-12\) with the aim of reducing the variability in data collection for international trauma registries. This template recommends using the newest version of AIS\(^11\), which is currently the AIS 2008 edition\(^1\). However, no reports have evaluated the inter-rater levels of agreement and reliability of the ISS and NISS based on this edition of the AIS. The aim of the present study was to estimate these parameters in a representative group of AIS-certified Norwegian trauma registry coders with a comparison against a reference standard.
Methods

Rater sampling

Study participants were recruited from a list of nineteen Norwegian trauma registry coders certified in the AIS system who were working in trauma registries or who were intended to code in hospitals in the process of establishing a registry. The list was cross-checked against a list of the Norwegian Better & Systematic Trauma Care Foundation network contact persons. Participating hospitals were compensated financially so that the raters could take time from their regular work to complete the coding for the study.

Case sampling

Patient records were selected from the trauma registry at Oslo University Hospital - Ullevål (OUH-U). OUH-U, the largest trauma centre in Norway, receives approximately 1400 trauma patients annually of which 34% have an ISS >15 (i.e., severe injury), and 44% have an NISS >15 according to AIS 2008. A sample of consecutive patients with NISS >15 who were directly admitted from the scene of injury were selected for the study. Exclusion criteria were asphyxia, drownings, and burns as the predominant injuries; hospital admission >24 hours after the injury; and patients declared dead before reaching the hospital and with no signs of life or response to initial resuscitation upon arrival to the Emergency Department (ED). Pre-hospital and ED charts, hospital admission notes, trauma anaesthesia records, surgery reports, physician’s progress notes, intensive care unit records, nurse reports, laboratory data, radiology reports, autopsy records (if applicable), and discharge summaries of each patient’s record were distributed. All direct patient identifiable information (e.g., name, address, and date of birth) and indirect patient identifiable data (e.g., the geographic location of the injury, age, and the date of injury) were replaced with fictitious data. The names of treating doctors,
treating hospital(s), signatures, departments, and dates of examinations, operations, and medical chart notes were deleted.

The determination of the sample size for ISS and NISS reliability studies depends upon a reasonable estimate of the reliability coefficient in the study population, the coefficient of the confidence interval (CI), and the maximum error (e.g., from a previous estimate)\textsuperscript{14, 15}. A sample of 50 cases has been suggested\textsuperscript{15} and was used in this study.

**Data variables and data collection**

The raters were asked to score all 50 cases according to an expanded version of the Utstein Trauma Template Dictionary\textsuperscript{11}, which contained 48 data variables, including AIS, ISS, and NISS. The raters were allowed to use either the AIS 2005 or AIS 2008. In cases where AIS 2005 ratings did not agree with AIS 2008 ratings, the AIS 2005 codes were converted to AIS 2008 codes using a mapping list\textsuperscript{16}. The ISS and NISS scores were calculated automatically by the study group.

Main outcome measures were completeness of injury coding, agreement in ISS and NISS scoring, and reliability of ISS and NISS scoring.

The raters reported their levels of experience and training in a questionnaire (Supplemental File 1). Two web-based data entry tools were used to collect data from the cases and the questionnaire.

**AIS, ISS, and NISS**

The AIS classification system is a consensus-derived, anatomically based, seven-digit global injury scoring system\textsuperscript{1}. The first six digits identify the injured body region (out of nine regions), the type of anatomic structure, and the specific anatomic structure; the seventh digit refers to an ordinal injury severity scale with categories ranging from 1 (‘minor injury’) to 6
(‘maximal injury’). From the AIS scores, an ISS value, a pragmatic quantitative summary measure of the overall severity of anatomic and functional damage, is calculated by summing the squares of the highest AIS severity codes in each of the three (out of the six) most severely injured ISS body regions\(^2\). In contrast, NISS, a revised version of the ISS, is calculated by summing the squares of the three most severe AIS injuries, regardless of ISS body region\(^3\). The ISS and NISS scales range from 1 to 75.

**Reference standard**

A reference panel developed a reference standard based on the AIS 2008 dictionary. The panel consisted of one trauma registrar who was an AAAM AIS certified course instructor and faculty member (MH), one trauma anaesthesiologist (NOS), and one PhD student (KGR), all experienced in AIS coding. Panellist coding disagreements were resolved by consensus or by consultation with clinical experts in the relevant surgical fields.

Each AIS code assigned by the raters was thoroughly examined against the reference standard to identify missing codes, coding errors, missing body regions, and other possible mistakes. In cases for which multiple skin injuries were assigned to a patient, the study investigators considered that coding one skin injury per severity level was sufficient. If a rater identified an obviously correct code that had not been identified by the reference panel, the code was added to the reference standard.

**Ethical considerations**

The Data Privacy Ombudsman for research at OUH considered this project exempt from license requirements, and the Regional Committee for Medical and Health Research Ethics (ref. no. 1.2009.1139, 2009/345-1) was informed but considered the project outside their
mandate for approval. Only raters who gave written informed consent were allowed to participate.

**Statistical methods**

Continuous data are presented as medians and ranges. Categorical data are presented as counts and percentages. The agreement and reliability between each rater and the reference standard (rater-against-reference) and the agreement and reliability between raters (inter-rater) were estimated for ISS and NISS. The same measures were estimated for the members of the reference standard group.

Inter-rater agreement was assessed using the Bland-Altman limits of agreement (LoA) method. This method compares the estimated variation in the data to a clinical evaluation of what is an acceptable variation in order for measurements to be considered “not different”. The LoA were calculated as the mean of the differences between the measurements of two raters ±1.96 × standard deviation, and will contain 95% of future measurements pairs in similar individuals, assuming a normal distribution of data. A smaller range between these two limits denotes better agreement. An extension of the LoA method to compare more than two pairs of measurements was used in this analysis. A clinically acceptable LoA range was set at ±9 units, equivalent to the increase in the derived ISS/NISS value when the severity of a single injury is increased from AIS 4 to 5.

Reliability is defined as the ratio of variation between measurements to the total variation of all the measurements it is intended to measure. Inter-rater reliability was estimated using intra-class correlation coefficient (ICC) statistics and corresponding 95% CI using a two-way mixed model with the absolute agreement index. ICC statistics give a number on a scale from 0-1, where 0 indicate agreement no better than expected by chance, and 1 indicates perfect agreement. The inter-rater ICC values were further analysed using a hierarchical
agglomerative cluster analysis (HAC) with complete linkage and ‘1-ICC’ as a distance measure for the accompanying dendrogram\textsuperscript{25}. HAC is a multivariate statistical method for partitioning values into optimally similar groups\textsuperscript{25}.

Univariate linear regression analysis was performed to evaluate the relationships between the ISS ICC and NISS ICC values as dependent variables and the number of cases coded during the rater’s entire career and years of coding experience as the independent variables. Statistical significance was assumed when \( P < 0.05 \).

Statistical analyses were performed using STATA/SE version 11.2 (StataCorp LP, College Station, TX, USA) and R version 2.11.1 (The R Foundation for Statistical Computing, Vienna, Austria)\textsuperscript{26}.

The Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were used in the drafting of this report\textsuperscript{21}.

**Results**

**Raters**

Of the 19 identified and invited raters, one declined to participate, and two never responded. Five that initially agreed to participate later withdrew due to resource constraints. One rater initially agreed to participate but did not respond to four e-mail reminders and never submitted the data. Ten raters answered a questionnaire and scored the 50 cases.

Three raters were clinically experienced registered nurses, five were specialist nurses (i.e., nurse anesthetists or critical care nurses), and two were physicians (anaesthesiologists). The median (range) coding experience was 3.5 (0,10) years, and the median (range) number of cases coded throughout the rater’s career was 275 (20,2000).

**Cases**
Of the 50 patient cases, 39 (78%) were male. The median (range) age was 45 years (7,83), and 90% were blunt traumas. Twelve (24%) of the patients died of their injuries. The median (range) ISS and NISS values in the reference standard were 22 (9,45) and 28 (16,75), respectively.

**AIS coding and injury identification**

The median (range) number of AIS codes assigned by the raters was 352 (275,459), compared to 382 codes assigned in the reference standard. Overall, the raters assigned 3561 AIS codes. Of the AIS codes, 2189 (61.5%) agreed with the reference standard. A total of 471 (13.3%) AIS codes contained errors (Table 1) but were still considered relevant and therefore included in the analyses. The raters recorded 392 injury codes that did not exist according to the reference standard, overlooked 1187 (31.1 %) injuries, and double-coded 509 injuries (e.g., coded more than one skin injury) (Table 1). The raters found 15 injuries that were not found by the reference panel. Six severity levels were changed in the reference standard, and nine injuries were removed.

**Agreement**

The LoAs between all raters and between each rater and the reference standard for ISS are shown in Figure 1a. The narrowest LoA range was from -8.12 to 10.48, and the widest range was from -35.98 to 35.22. Rater 1 disagreed markedly with all other raters and the reference standard. Rater 3 and 4 agreed the most with each other and the reference standard. The LoAs between all raters and between the raters and the reference standard for NISS are shown in Figure 1b. The narrowest LoA range was from -15.29 to 17.97, and the widest range was from -42.25 to 33.21. As for NISS, Rater 1 was least in agreement with all other raters.
and the reference standard. Rater 3 agreed the most with the reference, and raters 4 and 6 agreed the most with each other.

All LoAs, both for ISS and NISS, were wider than the predefined clinically acceptable LoA range of ±9.

**Reliability**

The joint ICC (range) for all raters against the reference standard was 0.51 (0.29,0.86) for ISS (Figure 3a) and 0.51 (0.27,0.78) for NISS (Figure 3b).

The joint ICC (range) for inter-rater reliability was 0.49 (0.19,0.85) for ISS, and 0.49 (0.16,0.82) for NISS. Hierarchical agglomerative clustering of the inter-rater estimated ISS ICC values revealed two subgroups, with raters 1 and 9 in the least agreement with the rest of the raters (Figure 4a). These two clusters showed relatively little agreement with one another. Inter-rater NISS ICC clustering also showed two subgroups (Figure 4b), confirming that raters 1 and 9 stood out as being the least in agreement with the other raters.

**Levels of agreement and reliability within the reference panel**

The narrowest ISS LoA for within-panel inter-rater agreement range was from -10.69 to 9.65, and the widest LoA was from -16.87 to 13.63. The narrowest NISS LoA range was from -13.44 to 16.24, and the widest range was from -25.75 to 18.83. All ranges were wider than the predefined clinically acceptable LoA range of ±9.

The joint ICC (range) for inter-rater reliability was 0.72 (0.64,0.79) for ISS and 0.68 (0.64,0.77) for NISS.

**Relationships between the ratings and the participants’ background and experience**
Univariate linear regression analyses indicated no statistically significant relationships between the rater-against-reference ICC and the total number of cases coded during the rater’s career for ISS (P=0.80) or NISS (P=0.45). Similarly, there was no significant relationship between rater-against-reference ICC and years of coding experience for ISS (P=0.39) or NISS (P=0.28).

**Discussion**

The anatomic injury scores assigned by ten AIS-certified trauma registry coders using AIS 2008 varied considerably, with less than two-thirds of the codes agreeing with a reference standard and with nearly one-third of injuries overlooked. This led to relatively low levels of agreement and reliability of injury severity scoring (ISS and NISS), and indicates that summative injury scoring using the AIS system is subject to large inter-rater variability and thus must be interpreted with great caution. We can probably assume that the AIS system will always be subject to some inter-rater variability. Therefore, the question is how large of a variation we can accept. In this study, we failed to find values within the acceptable LoA range of ±9 units of disagreement in ISS and NISS for the raters. Even more noteworthy was the amount of disagreement within the reference panel before panel consensus was reached.

Our study assessed the reliability of the instruments in a real-life situation. We used authentic, representative cases selected from consecutively admitted trauma patients, and the raters were instructed to independently score all 50 cases according to an expanded version of the Utstein Template. This study design permitted the evaluation of coding reliability and validity in a setting similar to that of real trauma registry work, including the registration and coding of multiple data variables, without us identifying specific injuries for the raters.

In contrast, MacKenzie et al. studied coding reliability by reviewing cases to identify a set of injuries that the majority of the raters had coded\(^7\), whereas Read-Allsopp asked her raters to
classify an artificially developed list of random injuries. Furthermore, we included a reference standard that enabled us to evaluate concurrent criterion-related validity, which involves a comparing of each rater’s assessment with that of a reference standard and allowed us to estimate the accuracy of injury identification and severity classification associated with the use of AIS codes. The few previous studies of agreement and/or reliability issues in the AIS system did not test these against a specifically developed consensus-derived reference standard. MacKenzie et al. identified one rater as the most accurate and consistent, and compared all other raters with this ‘standard’; however, these authors did not explain the specific criteria for accuracy and consistency that were used to choose the standard rater.

While the raters in our study had varying degrees of coding experience, they all had clinical experience in emergency medicine and trauma care, and all were certified in injury coding by the AAAM. They are thus considered a representative sample of those who currently code for the existing trauma registries in Norway and who will provide the data for a future national trauma registry. Thus, these results are a snapshot of the agreement and reliability that we may expect from Norwegian coders. We failed to find a relationship between rater-against-reference reliability and total number of cases coded or years of experience. However, we were unable to perform more detailed multivariate linear regression analyses because only ten raters were included. Therefore, we cannot rule out the possible influence of years of experience, cases coded in total, profession, or time spent on coding each case.

Decreasing the number of raters, i.e., performing centralized coding in a regional or national registry may reduce coding variability within the registry, but would not solve the problem of comparing data across trauma systems, nations, or between individual hospitals from different countries. Another way to improve the consistency of coding, especially with regard to the AIS coding, might be to have two coders screen each case. Alternatively, coders could
perform regular reviews of the accuracy of injury identification and AIS coding in a random set of cases. However, duplicate injury classification would increase costs in terms of human resources and training requirements, which is unaffordable and therefore unacceptable in resource-constrained settings\textsuperscript{29, 30}. This could result in AIS coding being performed only in hospitals with a zealous emphasis on trauma care (e.g., level I-II trauma centres in high-income countries)\textsuperscript{31}. Another possibility for injury coding could be to use the International Classification of Diseases Ninth Revision, Clinical Modification (ICD-9-CM)\textsuperscript{32} to map ICD-9-CM codes into AIS codes\textsuperscript{33}. However, most European hospitals use the ICD-10 edition, and even though an ICD-10-CM version is available\textsuperscript{34}, an ICD-10-CM-to-AIS 2008 mapping tool is, to our knowledge, not currently available. However, we anticipate that the ICD system will also be subject to inter-rater variability because the individuals assigning ICD codes will most likely face the same problems in identifying and coding injuries correctly according to radiology and surgery reports. Furthermore, in comparative studies of injury scores, the ICD-9-to-AIS-mapped scores generally did not perform as well as those based on directly coded AIS scores\textsuperscript{35-38}.

This study shows that, in our current setting, the use of the AIS methodology does not have adequate precision to function as a benchmarking tool. However, through targeted training, a more comprehensive AIS course, coding consensus processes, more time to code at each hospital, regular recertification, and properly designed databases, the accuracy, agreement and reliability of AIS scoring may increase to an acceptable level. Overall, these findings indicate the need to initiate quality improvement processes. Because a more adequate and precise injury-scoring alternative is not currently available, the trauma registry community may use these results to improve the scoring accuracy and precision. This is especially important for the introduction of a national inclusive trauma registry system.
Some of the variability between raters may also be caused by AIS scale limitations, or the ISS and NISS. AIS cannot be considered a biunique ordinal scale because the difference between AIS 1 and 2 is not the same as that between AIS 4 and 5, and the relationship between AIS score and mortality is not linear. Therefore, it may be difficult for raters to objectively judge injury severity. Another limitation of the AIS system that may cause rater variability may be that the AIS dictionary is very detailed and complex, with many specific coding rules, and cannot be properly understood without experience. The fact that the NISS summarizes the three most severe injuries regardless of body region may explain why the LoA ranges were wider and the ICC values were lower for NISS than for ISS.

Some limitations of the study are worth noting. The inclusion of only patients with a NISS >15 may have introduced a selection bias. However, because our focus was on severely injured patients, according to the inclusion criteria of the Utstein Trauma Template, this choice was considered valid. The process of assigning AIS codes may have been different for the reference group (fewer time restrictions, higher competence) than for the raters (more time restrictions, lower competence). The raters probably spent less time coding than the reference group due to greater time constraints, but constrained time frames are probably a more accurate reflection of the everyday setting for most raters. This may have affected the coding accuracy and reliability. The reliability might have been different for less injured patients, who are also more often treated and scored in lower-level trauma centres. Therefore, future studies should test the reliability of AIS scoring in patients with mild to moderate injuries.

Three panellists, all from the same institution, developed the reference standard. This may have introduced a bias as these panellists may have all adapted similar coding habits. However, all raters had attended courses with MH as an instructor, which should have reduced the variability between the reference panel and the raters.
Nine invited hospitals declined, did not respond to the invitation to participate in the study, or withdrew from the study. A lack of response from the invitees may have introduced a selection bias, although we have no reason to assume that the characteristics of the non-respondents differed greatly from those of the respondents.

This study did not test the reliability of the AIS scale as such; rather, it evaluated the accuracy of injury identification and the reliability of ISS and NISS. Therefore, we were not able to judge the actual reliability of AIS 2008.

Finally, because the medical charts contain information provided by several different health personnel, the injuries may have been described in several different ways in the same medical chart, in some cases. Injuries may have been described differently by radiologists and by surgeons, which may have caused some confusion in assigning the correct severity level.

Further studies may include an intra-rater reliability test and a new rater-against-reference standard reliability test. Further studies may also test the reliability of the AIS scale. Future studies should be designed to evaluate how differences in scoring between raters affect the utility of the ISS and NISS in outcome prediction models for trauma patients.

**Conclusions**

Anatomic injury scores assigned by AIS-certified trauma registry coders using AIS 2008 varied considerably in this study. This caused relatively low levels of agreement and reliability of injury severity scores for ISS and NISS and indicates that these scoring tools are overly rater dependent. ISS and NISS scores cannot be considered reliable classifiers for summarizing anatomic injury severity, and may not be appropriate for benchmarking trauma system performance.

**Conflict of interest statement**
KGR, MR, and AJK have received PhD funding from the Norwegian Air Ambulance Foundation. NOS has received Post Doc grants from the South-Eastern Norway Regional Health Authority. The other authors declare that they have no external financial or non-financial conflicts of interests related to this study.

Authors’ contributions

KGR, NOS, MH, MR, PAS, OR, AJK, and HML planned the study. KGR, MH, and MR selected and anonymised the medical records. KGR, MH, and AJK developed the web-based databases. KGR, MH, and NOS developed the reference standard, assisted by MR. KGR and MH investigated all injury codes. KGR and JR analysed the data. KGR wrote the first manuscript draft. All authors contributed to the interpretations of the results, helped to draft the manuscript, and approved the final version of the manuscript.

Acknowledgement

We thank senior lecturer J. Mary Jones PhD (Mathematics Department, Faculty of Natural Sciences, Keele University, Keele, UK) for assistance in planning the project. We acknowledge the Unit for Applied Clinical Research, Norwegian University of Science and Technology, Trondheim for designing the web-based database for the collection of clinical data from the trauma cases. We acknowledge The Centre for Information Technology Services, University of Oslo for providing a web-based tool for collecting questionnaire data. The SNLA funded the scholarships for the PhD students, Kjetil G. Ringdal MD, Marius Rehn MD, and Andreas J. Krüger MD. We also express deep gratitude to SNLA for financially compensating for the participants who took time from their hospital work to participate in the study. The study sponsor did not have any involvement in the study. We acknowledge the coding work performed by the participants from Finnmark Hospital, Hammerfest; Sørlandet
Hospital, Kristiansand; Nord-Trøndelag Hospital, Levanger; Nord-Trøndelag Hospital, Namsos; Oslo University Hospital, Ullevål; St. Olav’s University Hospital, Trondheim; Telemark Hospital, Skien; Vestre Viken Hospital, Asker and Bærum; and Østfold Hospital, Fredrikstad.
References


### Table 1. Description of the AIS codes assigned by each rater.

<table>
<thead>
<tr>
<th>Types of coding errors</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct AIS-coded injuries</td>
<td>159</td>
<td>197</td>
<td>261</td>
<td>228</td>
<td>227</td>
<td>249</td>
<td>209</td>
<td>262</td>
<td>171</td>
<td>226</td>
<td>2189</td>
</tr>
<tr>
<td>Completely correct</td>
<td>146</td>
<td>186</td>
<td>245</td>
<td>218</td>
<td>217</td>
<td>224</td>
<td>193</td>
<td>249</td>
<td>153</td>
<td>199</td>
<td>2030</td>
</tr>
<tr>
<td>Insignificantly different</td>
<td>13</td>
<td>11</td>
<td>16</td>
<td>10</td>
<td>10</td>
<td>25</td>
<td>16</td>
<td>13</td>
<td>18</td>
<td>27</td>
<td>159</td>
</tr>
<tr>
<td>Incorrect AIS-coded injuries</td>
<td>49</td>
<td>47</td>
<td>47</td>
<td>33</td>
<td>53</td>
<td>46</td>
<td>30</td>
<td>57</td>
<td>62</td>
<td>159</td>
<td>471</td>
</tr>
<tr>
<td>Severity too high</td>
<td>18</td>
<td>13</td>
<td>11</td>
<td>12</td>
<td>6</td>
<td>28</td>
<td>13</td>
<td>12</td>
<td>23</td>
<td>29</td>
<td>165</td>
</tr>
<tr>
<td>Severity too low</td>
<td>24</td>
<td>24</td>
<td>30</td>
<td>30</td>
<td>25</td>
<td>18</td>
<td>23</td>
<td>15</td>
<td>27</td>
<td>29</td>
<td>245</td>
</tr>
<tr>
<td>Incorrect organs or structures</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>Fundamental coding principle error</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>35</td>
</tr>
<tr>
<td>Non-existing injuries</td>
<td>45</td>
<td>50</td>
<td>59</td>
<td>23</td>
<td>18</td>
<td>71</td>
<td>26</td>
<td>18</td>
<td>36</td>
<td>46</td>
<td>392</td>
</tr>
<tr>
<td>Mistyped codes</td>
<td>3</td>
<td>20</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>66</td>
</tr>
<tr>
<td>Codes from different dictionary than that stated by the rater</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Injury not found in patient chart</td>
<td>42</td>
<td>30</td>
<td>47</td>
<td>18</td>
<td>17</td>
<td>61</td>
<td>18</td>
<td>17</td>
<td>31</td>
<td>41</td>
<td>322</td>
</tr>
<tr>
<td>Redundant injuries</td>
<td>45</td>
<td>50</td>
<td>59</td>
<td>23</td>
<td>18</td>
<td>71</td>
<td>26</td>
<td>18</td>
<td>36</td>
<td>46</td>
<td>509</td>
</tr>
<tr>
<td>Missed injuries</td>
<td>177</td>
<td>138</td>
<td>79</td>
<td>109</td>
<td>122</td>
<td>85</td>
<td>131</td>
<td>92</td>
<td>157</td>
<td>97</td>
<td>1187</td>
</tr>
<tr>
<td>AIS 1-2</td>
<td>118</td>
<td>86</td>
<td>51</td>
<td>68</td>
<td>73</td>
<td>43</td>
<td>74</td>
<td>62</td>
<td>103</td>
<td>65</td>
<td>743</td>
</tr>
<tr>
<td>AIS 3-5</td>
<td>59</td>
<td>52</td>
<td>28</td>
<td>41</td>
<td>49</td>
<td>42</td>
<td>57</td>
<td>30</td>
<td>54</td>
<td>32</td>
<td>444</td>
</tr>
</tbody>
</table>

AIS: Abbreviated Injury Scale.
Figure Legends

Figure 1. Bland-Altman limits of agreement for ISS and NISS scores from each rater against the reference standard and against each other.

This figure shows the agreement between raters and the agreement between each rater and the reference standard for ISS values (a) and NISS values (b), as expressed by the Bland-Altman 95% limits of agreement (LoA). The x-axis shows the LoAs, and the y-axis shows each pair of raters. The vertical broken line that crosses zero indicates perfect agreement, whereas the two dotted lines crossing ±9 indicates the clinically acceptable LoA.

Figure 2. Rater-against-reference reliability for ISS calculated by intraclass correlation coefficient.

The levels of reliability between each rater and the reference standard for ISS, determined by intraclass correlation coefficient (ICC) statistics and the corresponding 95% confidence intervals are (CI) depicted. ICC is based on a scale from 0-1, where 0 indicates agreement no better than that expected by chance, and 1 indicates perfect agreement.

Figure 3. Rater-against-reference reliability for NISS calculated by intraclass correlation coefficient.

The levels of reliability between each participant and the reference standard for NISS, determined by intraclass correlation coefficient (ICC) statistics and the corresponding 95% confidence intervals (CI) are depicted.

Figure 4. Hierarchical agglomerative clustering of intraclass correlation coefficient values for the inter-rater reliability of ISS (a) and NISS (b).
The dendrograms depict hierarchical agglomerative clustering. Similar elements are linked near the bottom of the graph (i.e., near 1), whereas dissimilar elements are linked higher on the graph (i.e., near 0).
Figure 1.
Figure 2.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>ICC Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1 vs Reference</td>
<td>0.292</td>
<td>0.013 - 0.527</td>
</tr>
<tr>
<td>R2 vs Reference</td>
<td>0.587</td>
<td>0.371 - 0.742</td>
</tr>
<tr>
<td>R3 vs Reference</td>
<td>0.855</td>
<td>0.756 - 0.915</td>
</tr>
<tr>
<td>R4 vs Reference</td>
<td>0.841</td>
<td>0.735 - 0.907</td>
</tr>
<tr>
<td>R5 vs Reference</td>
<td>0.727</td>
<td>0.513 - 0.847</td>
</tr>
<tr>
<td>R6 vs Reference</td>
<td>0.599</td>
<td>0.389 - 0.750</td>
</tr>
<tr>
<td>R7 vs Reference</td>
<td>0.733</td>
<td>0.572 - 0.839</td>
</tr>
<tr>
<td>R8 vs Reference</td>
<td>0.658</td>
<td>0.469 - 0.790</td>
</tr>
<tr>
<td>R9 vs Reference</td>
<td>0.443</td>
<td>0.191 - 0.640</td>
</tr>
<tr>
<td>R10 vs Reference</td>
<td>0.666</td>
<td>0.477 - 0.796</td>
</tr>
</tbody>
</table>
Figure 3.

<table>
<thead>
<tr>
<th>Pair</th>
<th>ICC Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1 vs Reference</td>
<td>0.318</td>
<td>0.053 - 0.544</td>
</tr>
<tr>
<td>R2 vs Reference</td>
<td>0.270</td>
<td>0.012 - 0.500</td>
</tr>
<tr>
<td>R3 vs Reference</td>
<td>0.763</td>
<td>0.609 - 0.860</td>
</tr>
<tr>
<td>R4 vs Reference</td>
<td>0.589</td>
<td>0.289 - 0.766</td>
</tr>
<tr>
<td>R5 vs Reference</td>
<td>0.344</td>
<td>0.069 - 0.570</td>
</tr>
<tr>
<td>R6 vs Reference</td>
<td>0.486</td>
<td>0.244 - 0.671</td>
</tr>
<tr>
<td>R7 vs Reference</td>
<td>0.630</td>
<td>0.275 - 0.806</td>
</tr>
<tr>
<td>R8 vs Reference</td>
<td>0.781</td>
<td>0.645 - 0.869</td>
</tr>
<tr>
<td>R9 vs Reference</td>
<td>0.528</td>
<td>0.299 - 0.701</td>
</tr>
<tr>
<td>R10 vs Reference</td>
<td>0.783</td>
<td>0.646 - 0.870</td>
</tr>
</tbody>
</table>
Figure 4.