ORIGINAL RESEARCH

Agreement Between Standard and ICD-10-Based Injury Severity Scores

Trine O Eskesen [ID][1]
Martin Sillesen[2,3]
Lars S Rasmussen [ID][1,3]
Jacob Steinmetz[1,3,4]

[1]Department of Anesthesia, Center of Head and Orthopedics, Rigshospitalet, Copenhagen, Denmark; [2]Department of Surgery and Transplantation, Rigshospitalet, Copenhagen, Denmark; [3]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; [4]Danish Air Ambulance, Aarhus, Denmark

Introduction: Injury Severity Score (ISS) is used to describe anatomical lesions. ISS is traditionally determined through medical record review (standard ISS), which requires specific training and may be time-consuming. An alternative way to obtain ISS is by use of ICD-9/10 injury diagnoses, and several conversion tools exist. We sought to evaluate the agreement between standard ISS and ISS obtained with two tools converting ICD-10 diagnoses.

Methods: Our cohort consisted of trauma patients ≥18 years admitted to Rigshospitalet between 1999 and 2016. The included patients had standard ISS recorded in the Trauma Audit and Research Network (TARN) database (ISS-TARN), and ICD-10 injury diagnoses for the trauma contact were recorded in the Danish National Patient Registry. We used the tools ICDPIC-R and ICD-AIS map to calculate ISS based on ICD-10 diagnoses. ICDPIC-R provided two ISSs: ISS-TQIP and ISS-NIS. The ICD-AIS map resulted in one ISS: ISS-map. The ISS-TARN was compared to the conversion tool ISSs using Bland-Altman plots. The agreement between ISS-TARN and the conversion tool ISSs for ISS above 15 was assessed using kappa statistics (κ).

Results: We included 4308 trauma patients. The median age was 44 years, 70% were male, and 92% had a blunt injury mechanism. The median ISS-TARN was 16 [IQR: 9–25], and the median conversion tool ISSs were 10 [2–25] (ISS-TQIP), 17 [5–26] (ISS-NIS), and 9 [4–16] (ISS-map). The Bland-Altman plots all showed increasing difference in ISS with increasing mean ISS. Bias ranged from −7.3 to 1.1 and limits of agreement ranged between −28.0 and 25.7. The agreement for ISS above 15 was fair to moderate (κ = 0.43 (ISS-TQIP), 0.44 (ISS-NIS), and 0.29 (ISS-map)).

Conclusion: Using ICDPIC-R or ICD-AIS map to determine ISS is feasible, but limits of agreement were unacceptably wide. The agreement between ISS-TARN and ICDPIC-R was moderate for ISS above 15.

Keywords: trauma, injury coding, ISS, ICD

Introduction

A description of injury severity is paramount in trauma research. Injury severity is known to be associated with trauma mortality[1] and is an important measure when comparing different trauma systems and trauma populations.

A widely used measure of anatomical injury is the Injury Severity Score (ISS).[1] To determine the ISS, a trauma patient's specific anatomical lesions are each given an Abbreviated Injury Scale (AIS) severity score from 1 to 6 as originally described by the American Medical Association (the Committee on Medical Aspects of Automotive Safety) and later revised multiple times by the Association for the Advancement of Automotive Medicine (AAAM).[2,3] The ISS is then derived from the AIS scores as "the

Correspondence: Trine O Eskesen
Department of Anesthesia, Center of Head and Orthopedics, Rigshospitalet, section 6011, Inge Lehmanns Vej 6, Copenhagen, DK-2100, Denmark
Tel +45 35 45 82 11
Email trine.grodum.eskesen.01@regionh.dk

201

sum of the squares of the highest AIS grade in each of the three most severely injured areas".[1] The standard method of determining the AIS scores and subsequent ISS is based on detailed review of the individual patient's medical record by a certified coding specialist. However, this requires specific training and may be time-consuming, and sometimes not possible when medical records cannot be retrieved.

An alternative to the standard medical record-based ISS is an ISS obtained from International Classification of Diseases (ICD) injury diagnoses. Here, a trauma patient's ICD injury diagnoses are converted into AIS severity scores and body regions from which an ISS can be calculated. In many countries, hospitals are required to record ICD diagnosis codes for all admitted patients, so the calculation of ISS from ICD injury diagnoses can save time and money. Several ICD-to-ISS conversion tools have been developed,[4–9] but most are made specifically for American ICD-Clinical Modification (CM) diagnosis codes (a modification of the original World Health Organization [WHO] ICD diagnosis codes), and they are not all generally available.

Recently, two conversion tools, the "Programs for Injury Categorization using the ICD and R statistical software" (ICDPIC-R) and the ICD-AIS map,[4,5] have been modified to handle the original 4-digit WHO ICD injury diagnoses instead of only the country-specific ICD injury diagnoses. Based on extensive and high-quality Danish health registries we aimed to evaluate these conversion tools using a large Danish trauma cohort.

The objective of this study was to evaluate the agreement between the standard ISS and the ISSs determined by two ICD-to-ISS conversion tools; ICDPIC-R and ICD-AIS map.

Materials and Methods
Ethics and Approvals
The handling of data in the present study complied with the current rules and regulations set forward by the data responsible institution in the Capital Region of Denmark; approval number P-2020-783. No consent or ethics committee was required for registry studies according to Danish law.

Study Design
This study was a retrospective cohort study comparing different methods of determining ISS for trauma patients. The standard method for determining ISS was by review

of medical records by a certified AIS coding specialist as mandated by the AAAM. This standard ISS (ISS-TARN) was compared with two alternative methods for determining ISS using the ICDPIC-R and the ICD-AIS map constructed to convert ICD-9/10 injury diagnoses to AIS severity scores for each body region and an ISS is then calculated subsequently. We used the Danish version of the ICD-10 injury diagnoses as input to the ICDPIC-R and ICD-AIS map. The Danish ICD-10 injury diagnoses consisted of up to 6 letters/digits, where the first and sixth letter were specific to Denmark. When these two letters were removed the Danish ICD-10 injury diagnoses corresponded to the 4-digit WHO ICD-10 injury diagnoses.

Study Population
The study population consisted of trauma patients ≥18 years of age admitted to the Trauma Center at Rigshospitalet (RH), Copenhagen, Denmark, between January 1999 and December 2016. RH hosts a level 1 trauma center with approximately 1000 trauma team activations per year. For patients to be included in this study, they had to be part of the RH cohort in the Trauma Audit and Research Network (TARN) database.[10] The TARN database includes patients of all ages who sustain injury resulting in hospital admission ≥3 days, critical care admission, transfer to a tertiary/specialist center or death within 30 days of hospital arrival. Isolated femoral neck or single pubic ramus fractures in patients ≥65 years and simple isolated injuries are excluded.[11] Patients identified from RH in the TARN database were then identified in the Danish National Patient Registry (DNPR) using the Danish personal identification number and date of trauma.[12] Patients were excluded if key data (personal identification number or trauma date) were missing. For patients with two admissions, both were kept in the dataset as the classification of the injury(ies) related to one trauma admission was considered not to be related to the other trauma admission.

Data Sources and Variables
Two data sources were used for this study; the TARN database (RH cohort) and the DNPR. From the TARN database we extracted information about demographics, mechanism of injury, date of injury, date of hospital arrival, date of hospital discharge, AIS severity score/body region, and standard ISS determined by a certified coding specialist (ISS-TARN). From 1999 to 2008 the TARN database used the AIS 1998 dictionary to determine AIS severity scores and body regions, and from 2008 to 2016 the AIS 2005 Revision 2008 dictionary was used. DNPR

supplied ICD-10 injury diagnoses from the hospital contact related to trauma. Relevant ICD-10 injury diagnoses were defined as ICD-10 diagnosis code S00-S99 (Danish version: DS00-DS99) and T00-T35 (Danish version: DT00-DT35).

Injury Severity Scoring

The AIS scores are anatomically based and derived by expert consensus. An AIS severity score describes a single injury on an ordinal scale ranging from 1 to 6, where 1 is "minor" injury and 6 is "maximal (currently untreatable)" injury.[3] Each AIS score is assigned to one of nine AIS chapters representing different body regions. However, for calculation of ISS the nine AIS chapters are reduced into six ISS body regions (head or neck, face, chest, abdominal or pelvic contents, extremities or pelvic girdle, external), and the ISS is the sum of the square of the highest AIS scores in the three most severely injured ISS body regions.[1] The ISS is an ordinal scale from 1 to 75. An ISS of 75 can be obtained either by three AIS scores of 5 or by just one AIS score of 6.[3] An ISS score below 15 is considered a minor traumatic injury, and a score above 15 is a major traumatic injury. It is not possible to obtain an ISS of exactly 15.

Conversions Tools

The ICDPIC is a conversion tool originally made for the statistical program STATA[13] but was in 2018 further developed to be used in the statistical program R and hence named ICDPIC-R (R package = icdpicr).[4,14] ICDPIC-R converts ICD-9 or ICD-10 injury diagnoses to AIS severity scores and body region by use of a data generated conversion algorithm. Here, an effect estimate for mortality for each ICD-10 injury diagnosis has been determined by logistic ridge regression. Then, using data-determined cut points for the effect estimates of the ICD-10 diagnosis code, each ICD-10 diagnosis code is assigned an AIS severity score.[15] There are two versions of the ICDPIC-R conversion algorithm, which differ in their underlying data reference, ie one uses the 2017 Trauma Quality Improvement Program (TQIP) data, and the other uses the 2016 National Inpatient Survey (NIS) trauma data. The resulting ISS using the TQIP version of the ICDPIC-R will be referred to as ISS-TQIP, and the ISS resulting from the NIS version will be called ISS-NIS. Initially the ICDPIC-R only accepted the American ICD-9/10-CM injury diagnoses, but it has recently been extended with a version 1.0.0. to also accept the WHO basic 4-digit ICD-10 injury diagnoses.

The ICD-AIS map is a conversion tool developed by a group of experts in ICD- and AIS coding appointed by the AAAM.[5] The ICD-AIS map consists of two translation maps. The first, the ICD-Serious Injury map, is a dichotomous map, which merely translates an ICD-9/10 injury diagnosis to AIS < 3 or AIS 3+. The second map, the ICD-ISS map, translates ICD-9/10 injury diagnoses to an AIS score, AIS chapter, and ISS body region, and this map can be used to calculate ISS. The ISS resulting from ICD-ISS map will be referred to as ISS-map. The ICD-AIS map is based on the AIS 2005 Revision 2008 Update dictionary.[3] The map accepts both the American ICD-9/10-CM injury diagnosis codes and the WHO ICD-9/10 4-digit injury diagnosis codes. In this study we have used the ICD-ISS map v2.0.

Statistical Analyses

Baseline data on the studied population were reported using median with interquartile range (IQR) for continuous variables and with frequencies and percentages for categorical variables. The agreement between the standard ISS (ISS-TARN) and the conversion tool ISSs (ISS-TQIP, ISS-NIS, and ISS-map) was assessed using Bland-Altman plots. We performed a visual evaluation of the Bland-Altman plots and calculated bias (mean difference) and limits of agreement (range in which 95% of the data points are to be found). We also performed Bland-Altman plots for only a subset of patients having a mean ISS above 15. In addition, we logarithmically (natural logarithm) transformed the ISS values and performed Bland-Altman plots of the log-transformed data. Bias and limits of agreement were then back-transformed by taking anti-log (the exponential function that is the inverse of the natural logarithm), which gave the ratio between the two ISS measurements.[16] A modified Bland-Altman plot was made with ISS-TARN on the x-axis. We further calculated the proportional bias for each conversion tool as ([conversion tool ISS − ISS-TARN]/mean ISS)*100%. The proportional bias was split into three groups; reasonable bias (≤ ±15%), moderate bias (±15–50%), and extreme bias (> ±50%). For each conversion tool and bias group the age distribution and proportion of patients with severe head injury (defined as head AIS ≥ 3) were calculated. In addition, we dichotomized ISS with cut-offs at 9, 15, and 25, respectively. The agreement between the ISS-TARN and the conversion tool ISSs for the dichotomized ISSs was investigated using kappa statistics. The evaluation of the resulting kappa coefficients was done according to Table 1 (slightly modified from[17]).

Table 1 Strength of Agreement Using Kappa Statistics

Value of κ	Strength of Agreement
< 0.20	Poor
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Good
0.81–1.00	Very good

Note: Adapted with permission from Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159. Biometrics © 1977 International Biometric Society.[18]

R Studio version 4.0.5. was used for the statistical analysis.

Results

In the TARN database we identified 9391 trauma admissions to RH between 1999 and 2016. Of these, 4313 fulfilled the following criteria: age ≥18 years, had an ISS-TARN, had a valid personal identification number, and had a date of trauma available. When matched with DNPR, 4308 had corresponding trauma dates and were included for analysis. A total of 13 (0.3%) patients had two trauma admissions in the study period. The included population had a median age of 44 years [IQR: 29.5–60] and 70% were male. Blunt mechanism of injury was described in 92% and 8% had penetrating injury. The median ISS-TARN was 16 [IQR: 9–25], and 54% were severely injured with an ISS > 15. The median number of unique ICD-10 injury diagnoses registered for the trauma contact was 4 [IQR: 2–6] (Table 2). The median conversion tool ISSs were 10 [2–25] (ISS-TQIP), 17 [5–26] (ISS-NIS), and 9 [4–16] (ISS-map) (Table 3).

Bland-Altman plots revealed an increasing ISS difference with increasing mean ISS (Figures 1A, 2A, and 3A). For ICD-NIS (Figure 2A) the bias was 1.1 (95% confidence interval [CI]: 0.7–1.4) and the limits of agreement were −23.5 to 25.7 (95% CIs: [−24.2 - −22.9]; [25.0–26.3]). For patients with mean ISS above 15 the data points were more evenly distributed around the x-axis, though, with a larger bias and a wider range of limits of agreement (Figures 1B, 2B and 3B). For ISS-NIS the bias was 4.0 (95% CI: 3.4–4.5) and limits of agreement were −24.0 to 31.9 (95% CI: [−25.0 - −23.0]; [30.9–32.9]) (Figure 2B).

A modified Bland-Altman plot was made with ISS-TARN on the x-axis (Figures 1C, 2C, and 3C). From

Table 2 Demographics and Injury Characteristics for the TARN Trauma Population

	N = 4308
Age, median [IQR] (years)	44 [29.5;60]
Male sex, n (%)	3019 (70)
Mechanism of injury, n (%)	
Blunt	3944 (92)
Vehicle incident/collision	2236 (57)
Fall, > 2 meters	768 (19)
Fall, < 2 meters	336 (8.5)
Blows/blast/crush	174 (4.4)
Burn	12 (0.3)
Other	418 (11)
Penetrating	362 (8)
Stabbing	257 (71)
Shooting	105 (29)
ISS, median [IQR]	16 [9;25]
> 15, n (%)	2317 (54)
No. of ICD-10 injury diagnoses, median [IQR]	4 [2;6]

Abbreviations: TARN, Trauma Audit and Research Network; ISS, Injury Severity Score; IQR; Interquartile Range; ICD-10, International Classification of Diseases 10th revision.

Table 3 Summary Statistics of the Four Different Methods for Determining ISS

	ISS-TARN	ISS-TQIP	ISS-NIS	ISS-Map
Median	16	10	17	9
IQR	9–25	2–25	5–26	4–16
< 15 (%)	46	54	43	75
> 15 (%)	54	46	57	25

Abbreviations: ISS, Injury Severity Score; TARN, Trauma Audit and Research Network; TQIP, Trauma Quality Improvement Program; NIS, National Inpatient Survey; IQR, Interquartile Range.

these figures it is evident that when ISS-TARN is low, the conversion tool ISS (ISS-TQIP, ISS-NIS, and ISS-map) is generally larger than ISS-TARN; however, when ISS-TARN is high, the conversion tool ISS is generally lower than ISS-TARN.
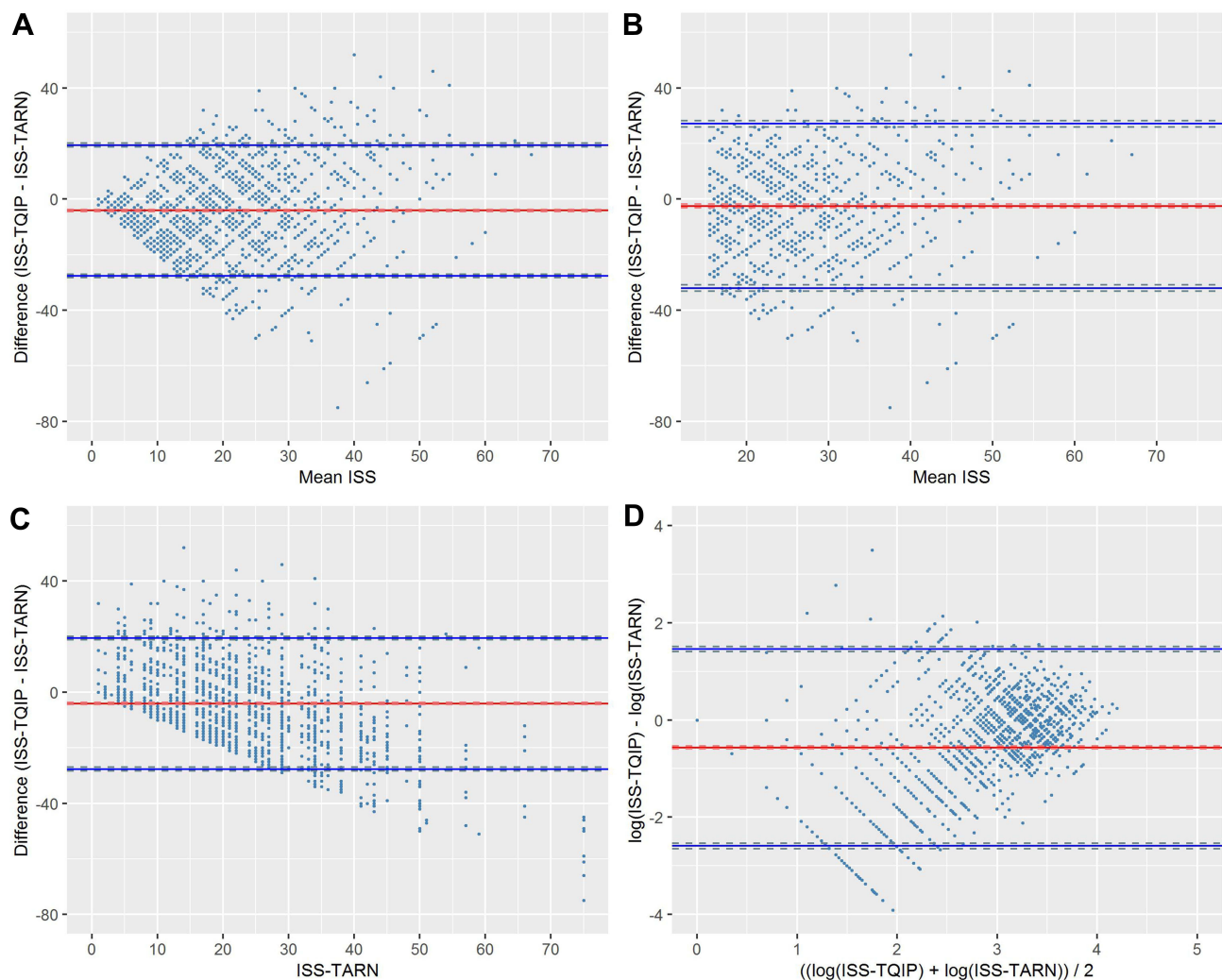
Figure 1 Bland-Altman plot comparing ISS-TQIP and ISS-TARN. (**A**) The difference between ISS-TQIP and ISS-TARN as a function of mean ISS ((ISS-TQIP + ISS-TARN)/2). Bias (red line): −4.1 (95% CI: −4.5 − 3.7). Limits of agreement (blue lines): −27.7 to 19.5 (95% CI: [−28.3 - −27.0]; [18.8–20.1]). (**B**) The difference between ISS-TQIP and ISS-TARN as a function of mean ISS ((ISS-TQIP + ISS-TARN)/2) for a subset of trauma patients with mean ISS > 15 (n=2012). Bias (red line): −2.5 (95% CI: −3.1 − −1.8). Limits of agreement (blue lines): −32.1 to 27.1 (95% CI: [−33.2 − −30.9]; [26.0–28.3]). (**C**) The difference between ISS-TQIP and ISS-TARN as a function of ISS-TARN. Bias (red line): −4.1 (95% CI: −4.5 − 3.7). Limits of agreement (blue lines): −27.7 to 19.5 (95% CI: [−28.3 - −27.0]; [18.8–20.1]). (**D**) The difference between log(ISS-TQIP) and log(ISS-TARN) as a function of mean log(ISS) ((log(ISS-TQIP) + log(ISS-TARN))/2). Proportional bias (red line): 0.57 (95% CI: 0.55–0.59), ie ISS-TQIP is on average 43% lower than ISS-TARN. Proportional limits of agreement (blue lines): 0.07–4.30 (95% CI: [0.07–0.08]; [4.08–4.54]), ie ISS-TQIP ranges from being 93% lower to 330% higher than ISS-TARN.

Abbreviations: ISS, Injury Severity Score; TQIP, Trauma Quality Improvement Program; TARN, Trauma Audit and Research Network.

The logarithmic transformation of the data and corresponding Bland-Altman plots did not yield normally distributed differences or a more even distribution of data points (Figures 1D, 2D, and 3D). For ISS-NIS the proportional bias (ratio between ISS-NIS and ISS-TARN) was 0.89 (95% CI: 0.87–0.92), ie ISS-NIS was on average 11% (95% CI: 8–13%) lower than ISS-TARN. The proportional limits of agreement for ISS-NIS ranged from 0.14 (95% CI: 0.14–0.15) to 5.61 (95% CI: 5.34–5.89), ie ISS-NIS ranged from being 86% lower to 461% higher than ISS-TARN (Figure 2D).

The proportional bias for the three conversion tool ISSs compared to ISS-TARN including the age distribution and proportion of patients with severe head injury is presented in Table 4.

The strength of agreement between ISS-TARN and the conversion tool ISSs (ISS-TQIP, ISS-NIS, and ISS-map) to categorize the injuries as mild (ISS < 15) or severe (ISS > 15) resulted in fair to moderate kappa coefficients of 0.43, 0.44, and 0.29, respectively. Additional dichotomization of ISS at cut-offs of 9 and 25 resulted in poor to moderate strengths of agreement (Table 5).
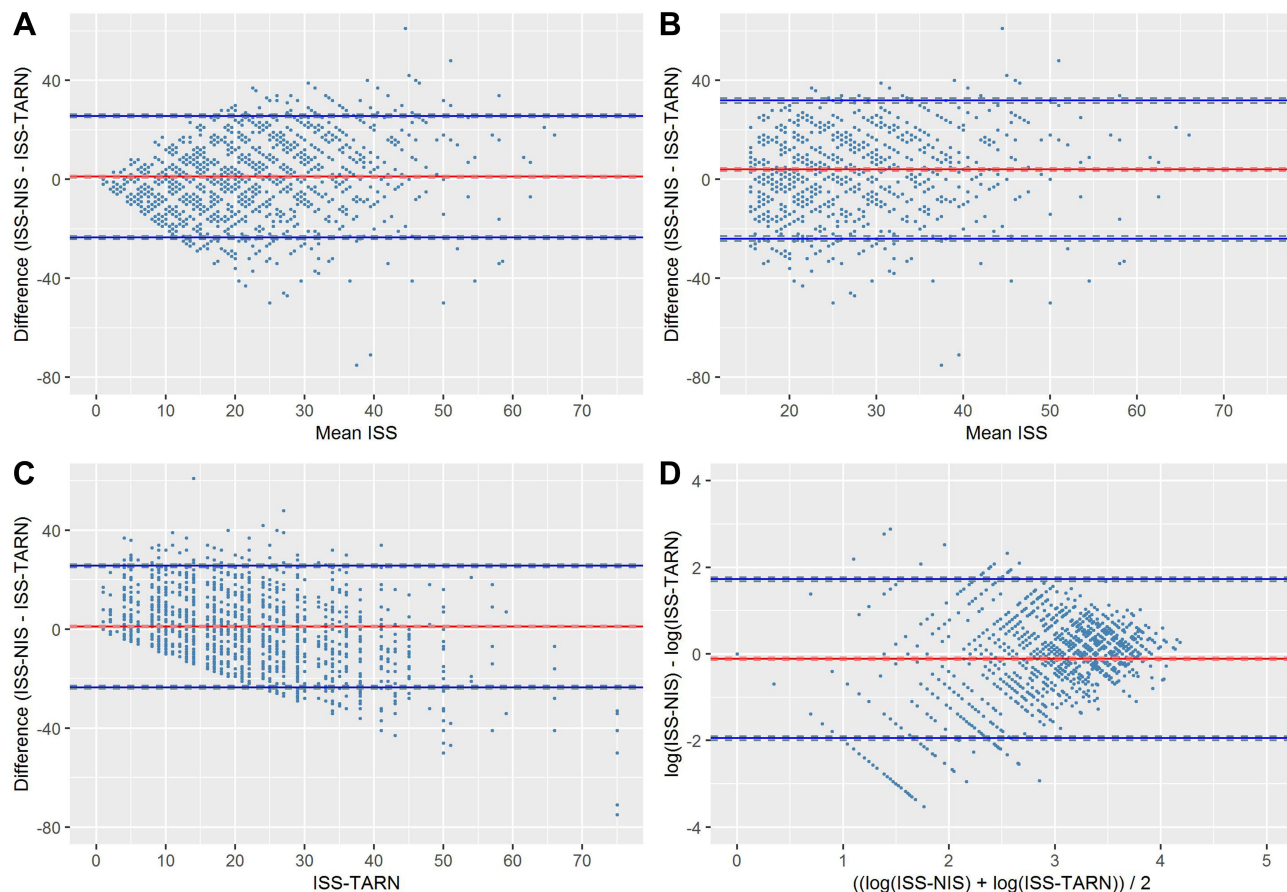
Figure 2 Bland-Altman plot comparing ISS-NIS and ISS-TARN. (**A**) The difference between ISS-NIS and ISS-TARN as a function of mean ISS ((ISS-NIS + ISS-TARN)/2). Bias (red line): 1.1 (95% CI: 0.7–1.4). Limits of agreement (blue lines): −23.5 to 25.7 (95% CI: [−24.2 - −22.9]; [25.0–26.3]). (**B**) The difference between ISS-NIS and ISS-TARN as a function of mean ISS ((ISS-NIS + ISS-TARN)/2) for a subset of trauma patients with mean ISS > 15 (n=2358). Bias (red line): 4.0 (95% CI: 3.4–4.5). Limits of agreement (blue lines): −24.0 to 31.9 (95% CI: [−25.0 – −23.0]; [30.9–32.9]). (**C**) The difference between ISS-NIS and ISS-TARN as a function of ISS-TARN. Bias (red line): 1.1 (95% CI: 0.7–1.4). Limits of agreement (blue lines): −23.5 to 25.7 (95% CI: [−24.2 - −22.9]; [25.0–26.3]). (**D**) The difference between log(ISS-NIS) and log(ISS-TARN) as a function of mean log(ISS) ((log(ISS-NIS) + log(ISS-TARN))/2). Proportional bias (red line): 0.89 (95% CI: 0.87–0.92), ie ISS-NIS is on average 11% lower than ISS-TARN. Proportional limits of agreement (blue lines): 0.14–5.61 (95% CI: [0.14–0.15]; [5.34–5.89]), ie ISS-NIS ranges from being 86% lower to 461% higher than ISS-TARN.
Abbreviations: ISS, Injury Severity Score; NIS, National Inpatient Survey; TARN, Trauma Audit and Research Network.

Discussion

In this study we assessed the agreement between standard ISS (ISS-TARN) and ISSs obtained by ICDPIC-R and ICD-AIS map when converting Danish ICD-10 injury diagnoses to ISS in an adult trauma population. We found that both the ICDPIC-R and the ICD-AIS map tended to overestimate the ISS when the ISS-TARN was low but underestimated the ISS when the ISS-TARN was high. The agreement between the ISSs derived by ICDPIC-R and ISS-TARN in distinguishing between minor and major injuries was moderate, whereas the agreement was only fair between ICD-AIS map derived ISS and ISS-TARN. Of the three ICD-to-ISS conversion methods, the ISS-NIS showed the highest agreement with ISS-TARN with bias close to 0, but with very wide limits of agreement.

The strength of this study lies in the use of a moderately large dataset from a well-defined region. Since only the ICD-10 version of injury diagnoses has been used, the data is consistent over a long time span. Since the ICD-10 codes are currently the primary ICD version in use, the results of this study may still be relevant for future studies. Another strength is that we have used the basic WHO ICD-10 injury diagnoses instead of country-specific diagnosis codes, which increases the generalizability of our findings well beyond the Danish trauma populations. Our study also has some limitation. First, the data originates from a single center. Second, the studied population includes only trauma patients fulfilling the TARN inclusion criteria. This has resulted in a population where more than 50% were severely injured,
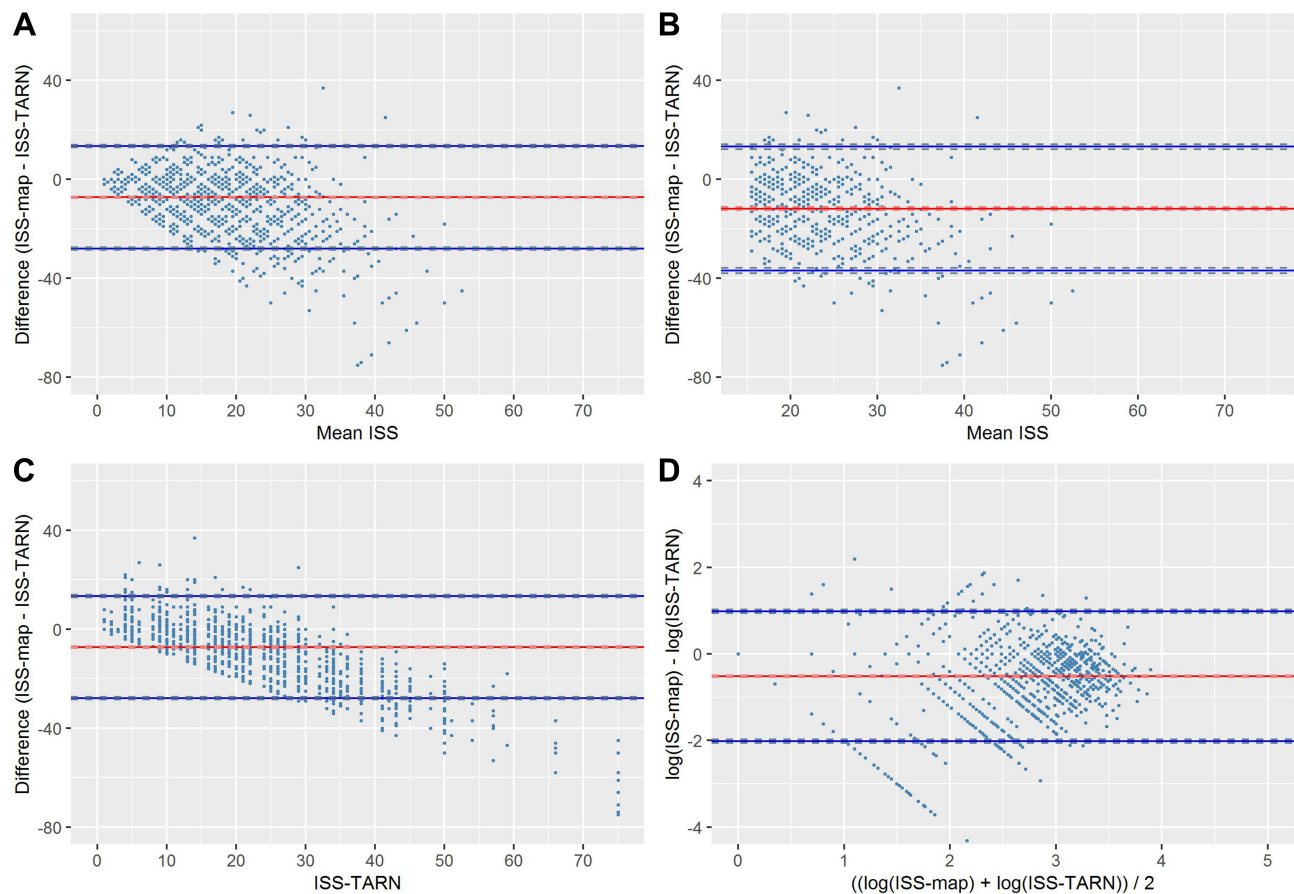
Figure 3 Bland-Altman plot comparing ISS-map and ISS-TARN. (**A**) The difference between ISS-map and ISS-TARN as a function of mean ISS ((ISS-map + ISS-TARN)/2). Bias (red line): −7.3 (95% CI: −7.6 − −7.0). Limits of agreement (blue lines): −28.0 to 13.4 (95% CI: [−28.5 − −27.4]; [12.9–14.0]). (**B**) The difference between ISS-map and ISS-TARN as a function of mean ISS ((ISS-map + ISS-TARN)/2) for a subset of trauma patients with mean ISS > 15 (n=1724). Bias (red line): −11.8 (95% CI: −12.4 − −11.2). Limits of agreement (blue lines): −36.8 to 13.2 (95% CI: [−37.9 − −35.8]; [12.2–14.2]). (**C**) The difference between ISS-map and ISS-TARN as a function of ISS-TARN. Bias (red line): −7.3 (95% CI: −7.6 − −7.0). Limits of agreement (blue lines): −28.0 to 13.4 (95% CI: [−28.5 − −27.4]; [12.9–14.0]). (**D**) The difference between log(ISS-map) and log(ISS-TARN) as a function of mean log(ISS) ((log(ISS-map) + log(ISS-TARN))/2). Proportional bias (red line): 0.60 (95% CI: 0.58–0.61), ie ISS-map is on average 40% lower than ISS-TARN. Proportional limits of agreement (blue lines): 0.13–2.68 (95% CI: [0.13–0.14]; [2.58–2.79]), ie ISS-map ranges from being 87% lower to 168% higher than ISS-TARN. **Abbreviations**: ISS, Injury Severity Score; TARN, Trauma Audit and Research Network.

which is much higher than the underlying RH trauma population in which 30% had sustained major trauma in 2018.[18] Third, the proportion of penetrating trauma was only 8%, thus the external validity for trauma populations with a higher proportion of penetrating trauma is low. Fourth, we have only included an adult trauma population, so our results are not valid for pediatric trauma.

The ICD diagnosis code system is developed by the WHO to provide a common language for classifying and monitoring virtually all possible diseases and is not developed specifically for describing injuries. The AIS, on the other hand, is developed with the exact purpose of describing anatomical injury with a level of detail that enables severity scoring and comparison. ICD and AIS codes are essentially different with differing levels of specificity, thus a specific injury in an individual patient may be

described quite differently by the ICD and the AIS coding systems.

The discrepancy between the standard ISS (ISS-TARN) and the ICD-10-based ISSs found in this study may be due to several factors. Programs for converting ICD-10 diagnoses to AIS scores and body region (and subsequently to ISS) do not recognize all ICD-10 injury diagnoses, which leaves some injuries unconverted and without the possibility to add to the final ISS. In general, this would most likely lead to lower ISS by conversion programs however, we found this to be true only for higher standard ISS values. A possible explanation for this difference between under- and overestimation may be found in the way ICD-10 codes are registered. In Denmark, ICD-10 codes are assigned by the attending physician, and it is more likely that a trauma patient with few injuries will

Table 4 Proportional Bias for the Three Conversion Tool ISSs Compared to ISS-TARN Split into Three Groups According to the Size of the Bias, Including the Age Distribution and Proportion of Patients with Severe Head Injury for Each Proportional Bias Group

	ISS-TQIP n = 4308	ISS-NIS n = 4308	ISS-Map n = 4308
Reasonable bias[a] (≤ ±15%), n (%)	717 (17)	852 (20)	934 (22)
Age, mean (sd) [years]	46.9 (19.2)	46.5 (19.0)	45.5 (18.9)
Head AIS3+, n (%)	382 (53)	380 (45)	206 (22)
Moderate bias[a] (±15–50%), n (%)	999 (23)	1257 (29)	1057 (25)
Age, mean (sd) [years]	48.2 (19.9)	47.3 (19.4)	45.9 (19.3)
Head AIS3+, n (%)	567 (57)	638 (51)	465 (44)
Extreme bias[a] (> ±50%), n (%)	2592 (60)	2199 (51)	2317 (54)
Age, mean (sd) [years]	45.3 (18.9)	45.5 (19.2)	46.7 (19.3)
Head AIS3+, n (%)	534 (21)	465 (21)	812 (35)

Notes: [a]The proportional bias was calculated as ([conversion tool ISS − ISS-TARN]/mean ISS)*100%, where conversion tool ISS is ISS-TQIP, ISS-NIS, and ISS-map, respectively, and mean ISS is the mean between ISS-TARN and the respective conversion tool ISSs.
Abbreviations: ISS, Injury Severity Score; TARN, Trauma Audit and Research Network; TQIP, Trauma Quality Improvement Program; NIS, National Inpatient Survey, sd; standard deviation, AIS, Abbreviated Injury Scale.

Table 5 Kappa Statistics for ISS Dichotomized at 9, 15, and 25, Respectively, for the Three Conversion Tool ISSs Compared to ISS-TARN

ISS Cut-Off	ISS-TQIP	ISS-NIS	ISS-Map
9	0.39	0.46	0.36
15	0.43	0.44	0.29
25	0.33	0.37	0.11

Abbreviations: TARN, Trauma Audit and Research Network; ISS, Injury Severity Score; TQIP, Trauma Quality Improvement Program; NIS, National Inpatient Survey.

have all injuries registered as ICD-10 codes, whereas a trauma patient with several injuries will not have all the injuries registered as this is more time-consuming. In contrast, the assignment of AIS scores is done by designated coding specialists with the primary task of identifying and scoring all injuries, thus it is more likely that all injuries will be recognized and coded.

Another aspect that might add to the discrepancy is the way the conversion tools are constructed. ICDPIC-R is based on a datadriven algorithm using data from two different databases (TQIP and NIS). These databases include all trauma patients from the data supplying centers; thus, a majority of the patients have sustained minor trauma (ISS < 15). This contrasts with the TARN population where 54% sustained major trauma (standard ISS > 15). The ICD-AIS map is, on the other hand, developed by an expert panel who has determined the

single ICD-to-AIS conversion by a multistep agreement process following general coding rules.[5] The map has primarily been made to fit the American ICD-10-CM injury code, thus when using the less specific 4-digit ICD-10 injury diagnoses as input the conversion may be less precise.

A third aspect to consider regarding the discrepancy is the AIS versions used. The AIS was originally established in 1971 but has since then gone through many revisions and updates.[3] The ICDPIC-R uses the original 1971 AIS version, whereas the ISS-TARN in this study is based on later versions (AIS 1998 and AIS 2005 Update 2008). The use of the oldest AIS version is a limitation of ICDPIC-R and may account for part of the disagreement between the standard ISS and ISS derived from ICDPIC-R.

In 1995, the DNPR changed from ICD-8 to ICD-10 for registering diagnoses, thus ICD-9 has never been in use in Denmark.[19] As a result, this study has only used ICD-10 injury diagnoses as input to the ICDPIC-R and ICD-AIS map. In contrast, most other studies evaluating the use of ICDPIC(-R) and ICD-AIS map are based on American data and thus use ICD-9-CM data, as the USA only changed to ICD-10-CM in 2015.[20] In particular, three large studies (>25,000 trauma patients) use ICDPIC(-R) to convert ICD-9-CM data to AIS/ ISS. The first study assessed the agreement between AIS determined by the coding specialist (standard) and AIS determined by ICDPIC for five separate body regions, and they found fair to good agreement (kappa coefficient ranging from

0.39 to 0.72).[21] In a second study the standard ISS was compared to ICDPIC-derived ISS. Like in our study, they found that the difference between the standard ISS and ICDPIC-derived ISS increased with increasing mean ISS. However, they found a good agreement (κ = 0.75) between the standard ISS and the ICDPIC-derived ISS when ISS was dichotomized at 15.[22] A third study split the ISS into four intervals (two intervals <15, two intervals >15) and found a weighted kappa of 0.74 (good agreement) when assessing the agreement between the standard ISS and the ICDPIC-derived ISS.[23] Only one study has used non-CM ICD-9 injury diagnoses as input to the ICDPIC, and found a kappa coefficient of 0.13 (poor agreement) when comparing standard ISS and ICDPIC-derived ISS split into three ISS-groups.[24]

A few studies have investigated the use of ICD-AIS map to determine ISS from ICD diagnoses. In a large study the ICD-AIS map was validated using both ICD-9-CM and ICD-10-CM injury diagnoses (vast majority of ICD-9-CM diagnoses). For the ICD-10-CM injury diagnoses perfect agreement between the standard ISS and the ICD-AIS map-derived ISS was found in 54%, increasing to 71% for exact ISS ± 5.[25] Another study used the ICD-serious injury map version of the ICD-AIS map to convert Finnish 4-digit ICD-10 injury codes to a dichotomous variable of AIS < 3 or AIS ≥ 3. They found a good agreement between the standard AIS and the ICD-serious injury map-derived AIS with a kappa coefficient of 0.70.[7]

The differences in ICD-version and the addition of clinical modification (CM) to American ICD injury diagnoses may account for some of the discrepancy between the findings in our and other studies. Also, the median ISS in our study was higher than in most other studies, where all but one had a median ISS below 15.[21–25]

Looking closer at the unconverted ICD-10 injury diagnoses in ICDPIC-R reveals some recurrent patterns. First, most basic 4-digit ICD-10 injury diagnoses concluding with "7" (describing "multiple injuries/fractures/ wounds/ etc.") are not converted by ICDPIC-R. Second, ICD-10 injury diagnoses describing "other/unspecified" body parts are not recognized by ICDPIC-R, although some of these ICD-10 injury diagnoses have very similar descriptions (but different diagnosis codes) in the ICDPIC-R conversion tables. This is likely because the ICDPIC-R conversion tables are originally made for ICD-10-CM diagnoses and then "shortened" to fit the basic 4-digit ICD-10 diagnoses. An improvement of the ICDPIC-R conversion tables based on a closer comparison between the 4-digit ICD-10 diagnoses and the current conversion tables may

decrease the number of unconverted diagnoses and improve the performance of ICDPIC-R.

From this study, as well as others, it is evident that using a conversion tool to determine ISS from ICD injury diagnoses is possible but has significant limitations. When possible the "hand coding" of ISS is preferred, as this allows single injuries to be described and coded with a higher precision and specificity than using a conversion tool. However, there are situations where a conversion tool may find its relevance for calculating ISS summary statistics. This could be in epidemiologic research where only ICD injury diagnoses, but not the ISS, are available for a trauma population and access to medical records is not possible.

Conclusions

Using ICDPIC-R and ICD-AIS map to convert 4-digit ICD-10 injury diagnoses to ISS is feasible, however, the limits of agreement were unacceptably wide. Of the three ICD-to-ISS conversion methods, the ISS-NIS showed the highest agreement with standard ISS and a moderate kappa for ISS above 15.

Abbreviations

AAAM, Association for the Advancement of Automotive Medicine; AIS, Abbreviated Injury Scale; DNPR, Danish National Patient Registry; ICD, International Classification of Diseases; ICDPIC-R, Programs for Injury Categorization using the ICD and R statistical software; IQR, Interquartile Range; ISS, Injury Severity Score; NIS, National Inpatient Sample; RH, Rigshospitalet; TARN, Trauma Audit and Research Network; TQIP, Trauma Quality Improvement Program; WHO, World Health Organization.

Data Sharing Statement

The data used in this research project cannot be shared due to the rules and regulations set forward by Statistics Denmark.

Ethics Approval and Informed Consent

No consent or ethics committee was required for registry studies according to Danish law.

Funding

have received financial support from The Danish Victims Fund (grant number 19-610-00046). The execution, content, and results of the materials are the sole responsibility of the authors. The analysis and viewpoints that have been made evident from the materials belong to the authors and do not necessarily reflect the views of The Council of The Danish Victims Fund.

Disclosure

The authors have no competing interests.

References

1. Baker SP, O'Neill B, Haddon W, Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14(3):187–196. doi:10.1097/00005373-197403000-00001
2. Committee on Medical Aspects of Automotive Safety A. Rating the severity of tissue damage. I. The abbreviated scale. *JAMA*. 1971;215 (2):277–280. doi:10.1001/jama.1971.03180150059012.
3. Gennarelli T, Wodzin EE. *Abbreviated Injury Scale (c) 2005 Update 2008*. Association for the Advancement of Automotive Medicine; 2016.
4. Clark DE, Black AW, Skavdahl DH, Hallagan LD. Open-access programs for injury categorization using ICD-9 or ICD-10. *Inj Epidemiol*. 2018;5(1). doi:10.1186/s40621-018-0149-8
5. Loftis KL, Price JP, Gillich PJ, et al. Development of an expert based ICD-9-CM and ICD-10-CM map to AIS 2005 update 2008. *Traffic Inj Prev*. 2016;17(S1):1–5. doi:10.1080/15389588.2016.1191069
6. Zonfrillo MR, Weaver AA, Gillich PJ, Price JP, Stitzel JD. New methodology for an expert-designed map from International Classification of Diseases (ICD) to Abbreviated Injury Scale (AIS) 3+ Severity Injury. *Traffic Inj Prev*. 2015;16(S2):S197–S200. doi:10.1080/15389588.2015.1054987
7. Airaksinen N, Nurmi-Lüthje I, Kröger H, Lüthje P. The ability of the ICD-AIS map to identify seriously injured patients in road traffic accidents—A study from Finland. *Traffic Inj Prev*. 2018;19 (8):819–824. doi:10.1080/15389588.2018.1520985
8. Haas B, Xiong W, Brennan-Barnes M, Gomez D, Nathens AB. Overcoming barriers to population-based injury research: development and validation of an ICD-10-to-AIS algorithm. *Can J Surg*. 2012;55(1):21–26. doi:10.1503/cjs.017510
9. Hartensuer R, Nikolov B, Franz D, Weimann A, Raschke M, Juhra C. Vergleich von ICD-10 und AIS mit der Entwicklung einer Methode zur automatisierten Umwandlung. *Z Orthop Unfall*. 2015;153 (06):607–612. doi:10.1055/s-0035-1546217
10. Trauma Audit and Research Network. Available from: www.tarn.ac. uk. Accessed December 10, 2021.
11. TARN Research Committee. The R in TARN. The Trauma Audit & Research Network; 2021.
12. Frank L. When an entire country is a cohort. *Science (80-)*. 2000;287 (5462):2398–2399. doi:10.1126/science.287.5462.2398
13. Clark DE, Osler TM, Hahn DR. ICDPIC: stata module to provide methods for translating International Classification of Diseases (Ninth Revision) diagnosis codes into standard injury categories and/or scores Statistical Software Components S457028. Boston College Department of Economy; 2009.
14. Black A, Clark D. icdpicr: "ICD" programs for injury categorization in R. R package version 1.0.0; 2021. Available from: https://cran. r-project.org/package=icdpicr. Accessed June 18, 2020.
15. Clark DE, Black AW. ICDPIC-R documentation. Available from: https://github.com/ablack3/icdpicr/commit/bf411d4dcb19ab1b8ddb f70e933c298665665cea. Accessed June 30, 2021.
16. Euser AM, Dekker FW, le Cessie S. A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. *J Clin Epidemiol*. 2008;61(10):978–982. doi:10.1016/j. jclinepi.2007.11.003
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159. doi:10.2307/2529310
18. Dansk Traumeregister Årsrapport 2018 (The Danish Trauma Registry Annual Report 2018); 2019. Available from: https://www.sundhed. dk/content/cms/19/97019_dansktraumeregister-aarsrapport_2018_offentligversion.pdf. Accessed December 10, 2021.
19. Sundhedsdatastyrelsen (The Danish Health Data Authority). Dokumentation af registre, Landspatientregisteret (Documentation of registries, The Danish National Patient Registry). Available from: https://www.esundhed.dk/Dokumentation/ DocumentationExtended?id=5. Accessed December 10, 2021.
20. Services USD of H& H. ICD-10 general equivalence mapping. Guidance for ICD-10 implementation. Available From: https://www. hhs.gov/guidance/document/icd-10-general-equivalence-mapping. Accessed December 10, 2021.
21. Greene NH, Kernic MA, Vavilala MS, Rivara FP. Validation of ICDPIC software injury severity scores using a large regional trauma registry. *INJ Prev*. 2015;21(5):325–330. doi:10.1136/injuryprev-2014-041524
22. Fleischman RJ, Mann NC, Dai M, et al. Validating the use of ICD-9 code mapping to generate injury severity scores. *J Trauma Nurs*. 2017;24(1):4–14. doi:10.1097/JTN.0000000000000255
23. Otto L, Wang A, Wheeler K, et al. Comparison of manual and computer assigned injury severity scores. *Inj Prev*. 2019;26 (4):330–333. doi:10.1136/injuryprev-2019-043224
24. Di Bartolomeo S, Tillati S, Valent F, Zanier L, Barbone F. ISS mapped from ICD-9-CM by a novel freeware versus traditional coding: a comparative study. *Scand J Trauma Resusc Emerg Med*. 2010;18(1):1–7. doi:10.1186/1757-7241-18-17
25. Glerum KM, Zonfrillo MR. Validation of an ICD-9-CM and ICD-10-CM map to AIS 2005 update 2008. *Inj Prev*. 2019;25(2):90–92. doi:10.1136/injuryprev-2017-042519