# Identifying individual-based injury patterns in multi-trauma road users by using an association rule mining method



Citation for the original published paper (version of record):
Fagerlind, H., Harvey, L., Humburg, P. et al (2022). Identifying individual-based injury patterns in multi-trauma road users by using an association
rule mining method. Accident Analysis and Prevention, 164.
http://dx.doi.org/10.1016/j.aap.2021.106479

Accident Analysis and Prevention 164 (2022) 106479

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

# Identifying individual-based injury patterns in multi-trauma road users by using an association rule mining method

Helen Fagerlind [a,b,d,*], Lara Harvey [a,b], Peter Humburg [a,c], Johan Davidsson [d], Julie Brown [a,f,g]

[a] Neuroscience Research Australia, Sydney, NSW 2031, Australia
[b] School of Population Health, University of New South Wales, Sydney, NSW 2052, Australia
[c] Stats Central, Mark Wainwright Analytical Centre, University of New South Wales, Sydney, NSW 2052, Australia
[d] Division of Vehicle Safety, Chalmers University of Technology, 412 96 Gothenburg, Sweden
[f] The George Institute for Global Health, Sydney, NSW 2042, Australia
[g] School of Medical Sciences, University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

*Keywords:*
Multiple injuries
Injury pattern
Road trauma
Road traffic crashes
Injury profiles
Injury outcome

ABSTRACT

In many road crashes the human body is exposed to high forces, commonly resulting in multiple injuries. This study of linked road crash data aimed to identify co-occurring injuries in multiple injured road users by using a novel application of a data mining technique commonly used in Market Basket Analysis. We expected that some injuries are statistically associated with each other and form Individual-Based Injury Patterns (IBIPs) and further that specific road users are associated with certain IBIPs. First, a new injury taxonomy was developed through a four-step process to allow the use of injury data recorded from either of the two major dictionaries used to document anatomical injury. Then data from the Swedish Traffic Accident Data Acquisition, which includes crash circumstances from the police and injury information from hospitals, was analysed for the years 2011 to 2017. The injury data was analysed using the Apriori algorithm to identify statistical association between injuries (IBIP). Each IBIP were then used as the outcome variable in logistic regression modelling to identify associations between specific road user types and IBIPs. A total of 48,544 individuals were included in the analysis of which 36,480 (75.1%) had a single injury category recorded and 12,064 (24.9%) were considered multiply injured. The data mining analysis identified 77 IBIPs in the multiply injured sample and 16 of these were associated with only one road user type. IBIPs and their relation to road user type are one step on the journey towards developing a tool to better understand and quantify injury severity and thereby improve the evidence-base supporting prioritisation of road safety countermeasures.

## 1. Introduction

Road traffic injuries are a global threat to people's health with 195 million road users affected during 2016 alone (Vos et al., 2017). In many road crashes the human body is exposed to high forces, commonly resulting in multiple injuries. To enable appropriate countermeasure prioritisation, there is therefore a need for injury outcome measurement tools that can adequately account for people with multiple injuries and differing patterns of co-occurring injuries among different road user types.

Most tools that take multiple injuries into account aim to predict the risk of mortality (Osler et al., 1997; Osler et al., 2019) and therefore may not adequately capture the impact of co-occurring injuries on long-term consequences for survivors. Whereas some injuries may induce a high threat to life, co-occurring injuries may generate a high risk of long-term disability. To-date it has been difficult to integrate this combined impact of individual injuries within a cluster of co-occurring injuries (Aharonson-Daniel et al., 2005; Aharonson-Daniel et al., 2003). Furthermore, despite a known increased risk of poorer functional and health status outcome at 12-months with increasing number of injuries sustained (Gabbe et al., 2014), many injury outcome studies only account for the primary documented injury (Halpin et al., 2009).

While there have been a few recent attempts to present injury outcome for road users in terms of patterns of injuries (Beck et al., 2017; Gabbe et al., 2015; Monárrez-Espino et al., 2018), these studies were not specifically focussed on the statistical association between different

* Corresponding author at: Neuroscience Research Australia, 139 Barker Street, Randwick Sydney, NSW 2031, Australia.
*E-mail addresses:* h.fagerlind@neura.edu.au (H. Fagerlind), l.harvey@neura.edu.au (L. Harvey), p.humburg@unsw.edu.au (P. Humburg), johan.davidsson@chalmers.se (J. Davidsson), jbrown@georgeinstitute.org.au (J. Brown).

https://doi.org/10.1016/j.aap.2021.106479
Received 3 February 2021; Received in revised form 5 October 2021; Accepted 5 November 2021
Available online 11 November 2021

types of co-occurring injuries. Yet there are some known associations between some specific types of injury that influence outcome. For example, it is well recognised that rib fractures increase the risk of pneumothorax and these types of injury commonly co-occur (Liman et al., 2003; Talbot et al., 2017). There may be many other associations between different injury types, forming Individual-Based Injury Patterns (IBIPs), that impact long term outcomes that have yet to be established. Knowing these would be useful to those developing injury countermeasures, as well as for acute treatment triage.

In this paper we describe a novel application of a data mining technique, which has been commonly used in Market Basket Analysis (Kotu and Deshpande, 2019) but not previously applied to injuries, to identify statistical associations between co-occurring injuries in multi-trauma road users.

## 2. Methods

### *2.1. Categorisation of injury codes to create a unified injury taxonomy*

A unified injury taxonomy was developed where specific injury codes from the Abbreviated Injury Scale 2005, update 2008 (AIS08) (AAAM, 2008) and the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) (WHO, 2016) were reduced into a smaller number of injury categories (than exist in either dictionary). There are two main reasons why this is necessary: to allow analysis based on data recorded by either dictionary to be compared; and to reduce the number of potential associations between injuries to a manageable number. The former is needed because AIS08 contains more detailed injury codes compared to the general WHO ICD-10 version and therefore, most injury codes do not translate one-to-one between the dictionaries. The latter is required because the number of possible associations between injuries (IBIPs) are related to how many unique injuries are entered into the computations. In previous work (Fagerlind et al., 2019) each individuals' injuries were collapsed into the most severe injury in each body region (according to AIS). It generated 835 unique patterns of injury of which 550 patterns only had one person assigned to an individual pattern. This amplification of possible injury patterns with small numbers of people within each pattern makes further analysis problematic.

In Appendix A we describe how the new injury taxonomy was developed through an extensive four-step process. The aim was to group injuries supported by published empirical results on the influence of each injury on four major groups of injury outcome measures: Survival/mortality; Acute injury; Impairment; or Disability (SAID). The resulting SAID taxonomy of 54 injury categories, each combining an anatomical location with the nature-of-injury, is presented in Table A1.

### *2.2. Association rule mining data analysis*

The Apriori Association Rule Mining algorithm (Agrawal et al., 1993) was used to identify co-occurring injury categories that exhibit statistically significant associations with each other in injured road users. The R package 'arules' was used to run the computations (Hahsler et al., 2005). In our analysis we named the resulting association rules Individual-Based Injury Patterns (IBIPs) which consists of one or more injury categories that co-occur with another injury category. To acquire the first IBIPs the association rule mining was restricted to individuals with two to five injury categories. Two injury categories are the minimum to be considered multiply injured. The maximum of five injury categories was selected through an iterative process as a trade-off between the gain in the number of individuals assigned a pattern and the number of patterns identified. All individuals, including those with more than five injury categories, with an IBIP injury combination were assigned that pattern. This may result in several patterns being assigned to the same individual. The IBIPs were selected to satisfy defined thresholds, described below, using the Support; Confidence; Lift; and Chi-square test statistic (Hahsler, 2015). Support is the frequency constraint and was set to include at least 30 individuals in each IBIP. In our analysis it represents the proportion of individuals that sustained a given combination of injury categories. Confidence is the conditional probability of sustaining one or more injury categories when another is present. The direction of the rule is disregarded, thereby treating any combination of injury categories as one. Confidence was selected to assign a pattern to at least 60% of individuals which resulted in a threshold of 0.2. In the mining process Chi-square test statistic and the Lift were calculated and IBIPs with p-value $< 0.05$ and Lift $> 1$ were considered robust patterns. Lift is the ratio of the observed frequencies and those expected if the injury categories are independent. Thus, Lift $>$ 1 indicate how much more likely the injury categories are to occur together then expected if they were statistically independent. The association rule output includes subsets of mined associations i.e. injury categories A; B; C; and D generates subsets of any combination of these categories, suggesting subsets are less severe. When this occurred, individuals were assigned the IBIP that included the highest number of injury categories for that specific combination. Age and gender profiles were computed for each IBIP as well as the proportion of individuals solely assigned one IBIP reported as 1-IBIP.

Bootstrapping was used to confirm the stability of the original IBIPs. One thousand random samples were drawn with replacement from the multiply injured dataset and then restricted to those with two to five injury categories. Every sample was used as input to compute sample IBIPs, as described for the original dataset. Finally, we examined the proportion of the samples in which each original IBIP was present.

### *2.3. Data source and selection*

Data from The Swedish Traffic Accident Data Acquisition (STRADA) was used to derive the IBIPs. The STRADA system is a census of all road crashes in Sweden, administrated by the Swedish Transport Agency. It includes crash circumstances from the police and injury information from hospitals. Injuries in STRADA are recorded by trained nurses at emergency departments according to the AIS08 dictionary (AAAM, 2008). Within the system, injuries are matched to crash records by date, time and location of the crash along with a personal identifier, creating a comprehensive road crash trauma source for Sweden (Howard and Linder, 2014). The injury data was used in the analysis of association between injuries (IBIP), and road user information was used in the analysis of association between IBIP and road user type. STRADA was queried from January 2011 to December 2017 for individuals aged less than 100 years old and injured in the road environment. Injuries were distributed into the SAID Taxonomy categories (Table A1) and injury categories of minor severity (Any body region superficial; Any body region sprain or strain; Finger fracture; and Tooth fracture) were excluded. Minor severity exclusions also included individuals with only a mild TBI concussion or Spine sprain or strain, or with only a combination of these. If any or both of these two injuries co-occurred with any other injury category they were included. Moderate concussions were included in the mild concussion category due to low numbers. After these exclusion, 49 injury categories remained and were considered in the analysis (Table A1).

Ethical approval for this study was obtained from the Regional Ethical Review Board in Gothenburg, Sweden (399-8) and Human Research Ethics Committee, University of New South Wales, Australia (HC180497). A waiver of consent was granted by the ethics committees.

### *2.4. Statistical analysis*

Descriptive statistical analyses were used to describe the data sample and demographics of road user type. We used a logistic regression model with each IBIP as the outcome variable and road user type as explanatory variables to identify specific road user types significantly associated to IBIPs. Road user type categories were: Car occupants; Cyclists in collision with another road user; Cyclists single; Moped riders (including passengers); Motorcycle riders (including passengers); and Pedestrians in collision with another road user. Truck and bus occupants were excluded due to low numbers. The reference road user category was selected according to which road user type experienced the highest proportion of each specific IBIP.

Data analyses were performed with R version 3.6.1 and RStudio (R Core Team, 2018; RStudio Team, 2016). Statistical significance was evaluated using $p < 0.05$.

## 3. Results

### 3.1. Overview of data

There were a total of 48,544 individuals in the trauma sample of which 36,480 (75.1%) had a single injury category recorded and 12,064 (24.9%) were considered multiply injured. Individuals with a single injury category may have sustained several injuries within the same SAID injury category (Table A1) but were not considered multiply injured in this study. Among the multiply injured, most individuals (61.4%) had two injury categories recorded.

In Fig. 1 the proportion of people sustaining the SAID injury categories (Table A1) in single category and multiple category injury samples are presented by body region. The proportion of each injury category within the multi-trauma sample is higher in almost all injury categories compared to the proportion present in the single category sample. Multiple rib fracture is the most common injury category in multiple injured people and have the largest difference in proportion between the samples. Radius fracture distal or shaft have the highest proportion of individuals in the single injury category sample but still have a higher proportion of affected individuals in the multiple injury sample.

### 3.2. Identification of individual-based injury patterns (IBIPs)

Of the 12,064 individuals considered multiply injured, 11,124 (92.2%) had 2–5 injury categories from which the data mining identified 69 injury patterns (Table 1, IBIP1–IBIP69). The stability of the resulting injury patterns was confirmed through bootstrapping. All patterns (IBIP1-IBIP69) were present in at least 481 bootstrap samples with 57 patterns (82.6%) present in over 900 samples. The bootstrapping generated another 119 patterns, but these were only present in less than 483 of the bootstrap samples.

Injury categories in the abdominal and thorax region as well as the head region created many similar combinations of IBIPs containing three injury categories. To decrease the number of assigned patterns for people with these combinations three IBIPs (all present in the bootstrap samples) containing four injury categories were added to the list of patterns (Table 1, IBIP70 – IBIP72). Finally, all individuals (>2 injury categories) that were not assigned any of the patterns 1 to 72 were run through the algorithm again which added five further patterns, confirmed by the bootstrapping, to the list (Table 1, IBIP73–IBIP77).

After the additional patterns were assigned to individuals with these injury combinations, 60.2% ($n = 7268$) of individuals with multiple injury categories were assigned a pattern. Of those people not assigned a pattern ($n = 4796$), 16.6% sustained more than two injury categories. Most IBIPs included injury categories from more than one body region, 37 patterns involved two, and 7 patterns involved three body regions. The remaining 33 IBIPs involved only one body region of which the most frequent ($n = 12$) was moderate to severe TBI injuries (TBI moderate+) in the head body region. Table 1 shows that the TBI injury categories: Focal brain injury; Subdural Haemorrhage (SDH); Subarachnoid Haemorrhage (SAH); and Skull fractures commonly co-occur. IBIPs with these four injuries in different combinations (Table 1, IBIP no. 2, 3, 6, 7, 9, 10, 11, 13, 14, 15, 70) are supersets or subsets of each other meaning an individual can only be assigned one of these patterns. There are 759 individuals with these TBI patterns of which 67.7% ($n = 514$) sustained a skull fracture.

### 3.3. Relationship between IBIPs and road user type

Knowing which injuries are related to certain types of road users is important when countermeasures are to be prioritised. Using the IBIP for multiply injured individuals could extend the precision of measures to be implemented. Table 2 reports the age and gender profile by road user type for multiply injured individuals. Car occupants and single cyclists are the most common road users. The median age is lowest for moped rider and highest for pedestrians. Males are predominant in all road user types except pedestrians.

In 16 of the 77 logistic regression models one type of road user was solely associated to the IBIP, with all other road users having significantly lower odds of sustaining these patterns (Table 3). Car occupants were associated with 5 injury patterns, notably thoracic or lumbar spine fractures co-occurring with cervical spinal fractures (IBIP41) or multiple rib fractures (IBIP61); and multiple rib fractures co-occurring with other lung injuries (IBIP49). Car occupants were also associated with sprain or strains in the spinal region (IBIP30, IBIP40). In contrast, single cyclists were associated with 3 IBIPs; involving upper limb fractures (IBIP17, IBIP31); or facial fractures with concussion (IBIP63). Motorcyclists were associated with 2 patterns predominantly involving the thoracic cage region and lungs (IBIP36) or internal abdominal injuries with rib and clavicle or scapular fractures (IBIP42). Pedestrians were associated with 5 patterns. Notably 4 of these patterns involved the pelvic region (IBIP8, IBIP35, IBIP56, IBIP77) and two patterns involved skull fracture co-occurring with intracranial brain injury or facial fracture (IBIP10, IBIP52).

Car occupants and single cyclist tend to have only one pattern while motorcyclists and pedestrians have patterns that are likely to appear in combination with other patterns. Cyclists in collisions with another road user and moped riders did not have any patterns significantly associated only to them. IBIPs that were associated with more than one road user type are reported in Appendix B. Table B.1.

## 4. Discussion

In this paper, a novel application of a data mining technique was used to identify IBIPs of associated co-occurring injury categories and some clear distinctions in IBIPs between different road user types were demonstrated. These results are useful in understanding how injury severity and outcomes can vary and might impact the prioritisation of crash countermeasures. The fact that some similar patterns are strongly associated to specific road users while others appear across all road users also reveal important differences for the development of interventions and design of protective equipment. Overall, this new understanding can add precision in prioritising and targeting countermeasures for regulators and road administrations.

Importantly in the approach taken in this work, all severity levels of trauma were considered except for minor injuries. By including moderate severity injuries, this approach allows disability outcomes to be better incorporated in evidence-based support for safety measure strategies. This is important given the major contribution of disability to poorer health outcomes (Gabbe et al., 2015; Polinder et al., 2015; Tournier et al., 2014), and the historical neglect of disability in setting priorities for crash injury countermeasures. The IBIPs generated in this work also provide a finer distinction of which injury categories are significantly associated with each other than could be obtained from analyses using only one primary diagnosis.

The novel approach taken in this work also addressed inherent differences between ICD-10 and AIS08 injury codes. In the present study (Appendix A), empirical data was used to rank the severity of injuries and collapse them into injury categories rather than grouping injuries by predefined severities provided in the AIS08 dictionary. In validating a

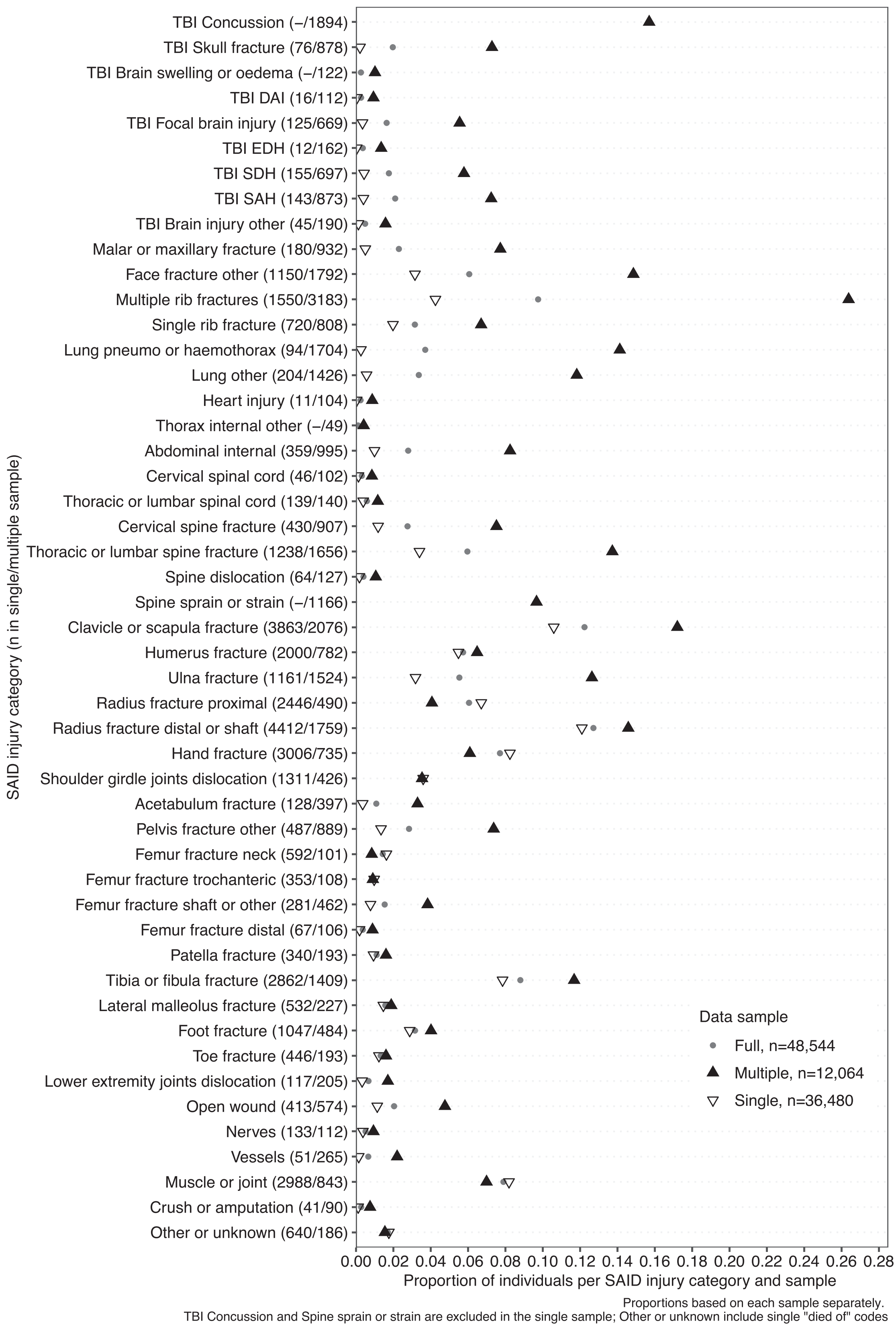


**Fig. 1.** The proportion of individuals sustaining a SAID injury category (by body region) in each of the sample: single injury category; multiple injury category; or in the full sample.

**Table 1**
Identified associated co-occurring injury categories presented as Individual-Based Injury Patterns (IBIPs) and ordered in descending order of Lift.

| IBIP | Individual-Based Injury Pattern (IBIP) | n | Lift | 1-IBIP [%] | Median age [IQR] | Males [%] |
|---|---|---|---|---|---|---|
| ibip0 | None assigned | 4796 | – | – | 48 [28, 62] | 61.8 |
| ibip1 | TBI EDH; TBI Skull fracture | 108 | 11.9 | 28.7 | 29 [19, 50] | 71.3 |
| ibip2 | TBI Focal brain injury; TBI SDH; TBI Skull fracture | 48 | 10.4 | 41.7 | 51 [28, 67] | 60.4 |
| ibip3 | TBI Focal brain injury; TBI SDH; TBI SAH | 34 | 9.3 | 58.8 | 58 [40, 71] | 73.5 |
| ibip4 | Cervical spine fracture; Spine dislocation | 67 | 9.1 | 58.2 | 47 [29, 64] | 77.6 |
| ibip5 | Foot fracture; Toe fracture | 64 | 8.8 | 70.3 | 40 [27, 61] | 76.6 |
| ibip6 | TBI SDH; TBI SAH; TBI Skull fracture | 77 | 8.7 | 53.2 | 59 [37, 72] | 61.0 |
| ibip7 | TBI Focal brain injury; TBI SAH; TBI Skull fracture | 63 | 8.0 | 31.7 | 49 [34, 59] | 73.0 |
| ibip8 | Acetabulum fracture; Pelvis fracture other | 174 | 7.1 | 47.7 | 59 [34, 72] | 56.3 |
| ibip9 | TBI SDH; TBI Skull fracture | 84 | 6.5 | 46.4 | 44 [24, 61] | 66.7 |
| ibip10 | TBI Focal brain injury; TBI Skull fracture | 102 | 6.4 | 39.2 | 39 [23, 63] | 70.6 |
| ibip11 | TBI SDH; TBI SAH | 98 | 6.3 | 46.9 | 63 [47, 73] | 73.5 |
| ibip12 | Humerus fracture; Shoulder girdle joints dislocation | 140 | 5.8 | 90.7 | 51 [41, 64] | 60.0 |
| ibip13 | TBI Focal brain injury; TBI SAH | 82 | 5.5 | 41.5 | 52 [33, 66] | 59.8 |
| ibip14 | TBI Focal brain injury; TBI SDH | 31 | 5.5 | 67.7 | 45 [22, 64] | 87.1 |
| ibip15 | TBI SAH; TBI Skull fracture | 60 | 5.5 | 60.0 | 53 [25, 66] | 70.0 |
| ibip16 | Malar or maxillary fracture; Face fracture other | 401 | 4.9 | 87.0 | 48 [32, 61] | 70.8 |
| ibip17 | Radius fracture distal or shaft; Ulna fracture | 1016 | 4.7 | 86.4 | 24 [12, 55] | 57.3 |
| ibip18 | Lower extremity joints dislocation; Tibia or fibula fracture | 104 | 4.7 | 62.5 | 50 [34, 62] | 43.3 |
| ibip19 | Abdominal internal; Vessels | 117 | 4.2 | 11.1 | 40 [29, 59] | 68.4 |
| ibip20 | Malar or maxillary fracture; Face fracture other; TBI Concussion | 104 | 3.8 | 79.8 | 52 [37, 61] | 66.3 |
| ibip21 | Malar or maxillary fracture; Face fracture other; TBI SAH | 87 | 3.6 | 18.4 | 59 [48, 68] | 67.8 |
| ibip22 | Lateral malleolous fracture; Tibia or fibula fracture | 81 | 3.5 | 69.1 | 55 [29, 66] | 46.9 |
| ibip23 | Malar or maxillary fracture; Face fracture other; TBI Skull fracture | 105 | 3.5 | 11.4 | 53 [30, 64] | 72.4 |
| ibip24 | Multible rib fractures; Lung pneumo or haemothorax; Clavicle or scapula fracture | 233 | 3.4 | 70.4 | 55 [45, 65] | 81.1 |
| ibip25 | Lung other; Multible rib fractures; Lung pneumo or haemothorax; Clavicle or scapula fracture | 143 | 3.3 | 28.7 | 48 [34, 58] | 79.0 |
| ibip26 | Cervical spine fracture; Thoracic or lumbar spine fracture; Multible rib fractures | 145 | 3.1 | 18.6 | 55 [46, 69] | 69.7 |

**Table 1** (*continued*)

| IBIP | Individual-Based Injury Pattern (IBIP) | n | Lift | 1-IBIP [%] | Median age [IQR] | Males [%] |
|---|---|---|---|---|---|---|
| ibip27 | Multible rib fractures; Lung pneumo or haemothorax | 309 | 3.0 | 91.9 | 57 [46, 67] | 75.7 |
| ibip28 | Foot fracture; Tibia or fibula fracture | 170 | 3.0 | 48.2 | 46 [29, 60] | 67.1 |
| ibip29 | Thoracic or lumbar spine fracture; Multible rib fractures; Lung pneumo or haemothorax | 187 | 3.0 | 23.0 | 49 [38, 64] | 73.3 |
| ibip30 | Other or unknown; Spine sprain or strain | 46 | 3.0 | 82.6 | 48 [25, 60] | 67.4 |
| ibip31 | Radius fracture proximal; Ulna fracture | 187 | 2.9 | 80.7 | 23 [11, 52] | 62.0 |
| ibip32 | Femur fracture shaft or other; Tibia or fibula fracture | 175 | 2.9 | 30.3 | 41 [26, 63] | 64.6 |
| ibip33 | Abdominal internal; Multible rib fractures; Lung pneumo or haemothorax | 140 | 2.9 | 34.3 | 52 [38, 63] | 70.0 |
| ibip34 | Abdominal internal; Lung other; Lung pneumo or haemothorax | 51 | 2.9 | 25.5 | 23 [18, 25] | 88.2 |
| ibip35 | Pelvis fracture other; Multible rib fractures; Lung pneumo or haemothorax | 187 | 2.9 | 5.3 | 53 [33, 67] | 66.3 |
| ibip36 | Lung other; Multible rib fractures; Clavicle or scapula fracture | 93 | 2.8 | 48.4 | 46 [33, 59] | 78.5 |
| ibip37 | Cervical spine fracture; Multible rib fractures; Lung pneumo or haemothorax | 122 | 2.8 | 8.2 | 53 [43, 68] | 72.1 |
| ibip38 | Abdominal internal; Lung other | 78 | 2.8 | 61.5 | 25 [18, 41] | 82.1 |
| ibip39 | Lung other; Multible rib fractures; Lung pneumo or haemothorax | 140 | 2.8 | 48.6 | 48 [29, 63] | 78.6 |
| ibip40 | Muscle or joint; Spine sprain or strain | 205 | 2.7 | 97.1 | 34 [21, 49] | 52.7 |
| ibip41 | Cervical spine fracture; Thoracic or lumbar spine fracture | 190 | 2.6 | 69.5 | 49 [31, 64] | 72.6 |
| ibip42 | Abdominal internal; Multible rib fractures; Clavicle or scapula fracture | 120 | 2.6 | 12.5 | 44 [30, 57] | 80.0 |
| ibip43 | Lung other; Lung pneumo or haemothorax | 75 | 2.5 | 66.7 | 23 [20, 37] | 76.0 |
| ibip44 | Thoracic or lumbar spine fracture; Lung other; Lung pneumo or haemothorax | 167 | 2.5 | 1.8 | 41 [25, 58] | 75.4 |
| ibip45 | Tibia or fibula fracture; Multible rib fractures; Lung pneumo or haemothorax | 134 | 2.2 | 9.7 | 54 [41, 66] | 59.7 |
| ibip46 | Patella fracture; Tibia or fibula fracture | 65 | 2.2 | 35.4 | 45 [29, 69] | 55.4 |
| ibip47 | Thoracic or lumbar spine fracture; Lung other; Multible rib fractures | 224 | 2.2 | 11.6 | 47 [31, 60] | 76.3 |
| ibip48 | Abdominal internal; Lung pneumo or haemothorax | 41 | 2.2 | 31.7 | 29 [20, 44] | 70.7 |
| ibip49 | Lung other; Multible rib fractures | 195 | 2.1 | 82.6 | 53 [40, 68] | 73.3 |

(*continued on next page*)

**Table 1** (*continued*)

| IBIP | Individual-Based Injury Pattern (IBIP) | n | Lift | 1-IBIP [%] | Median age [IQR] | Males [%] |
|---|---|---|---|---|---|---|
| ibip50 | Abdominal internal; Lung other; Multible rib fractures | 103 | 2.1 | 37.9 | 41 [26, 57] | 74.8 |
| ibip51 | Abdominal internal; Multible rib fractures | 112 | 2.0 | 83.9 | 56 [36, 64] | 71.4 |
| ibip52 | Face fracture other; TBI Skull fracture | 129 | 1.9 | 31.8 | 39 [24, 57] | 72.9 |
| ibip53 | Lung pneumo or haemothorax; Clavicle or scapula fracture | 64 | 1.9 | 59.4 | 42 [27, 56] | 84.4 |
| ibip54 | Thoracic or lumbar spine fracture; Multible rib fractures; Clavicle or scapula fracture | 62 | 1.8 | 30.6 | 54 [37, 64] | 79.0 |
| ibip55 | Multible rib fractures; Clavicle or scapula fracture | 326 | 1.8 | 89.0 | 55 [47, 65] | 78.8 |
| ibip56 | Pelvis fracture other; Multible rib fractures; Clavicle or scapula fracture | 106 | 1.8 | 16.0 | 60 [38, 73] | 57.5 |
| ibip57 | Abdominal internal; Thoracic or lumbar spine fracture; Multible rib fractures | 174 | 1.8 | 9.8 | 48 [33, 61] | 70.7 |
| ibip58 | Face fracture other; TBI SAH; TBI Skull fracture | 104 | 1.7 | 14.4 | 50 [31, 62] | 72.1 |
| ibip59 | Face fracture other; TBI SAH | 91 | 1.7 | 44.0 | 60 [47, 67] | 73.6 |
| ibip60 | Lung other; Lung pneumo or haemothorax; Clavicle or scapula fracture | 31 | 1.6 | 19.4 | 26 [19, 42] | 87.1 |
| ibip61 | Thoracic or lumbar spine fracture; Multible rib fractures | 228 | 1.5 | 84.6 | 60 [47, 71] | 61.0 |
| ibip62 | TBI Concussion; Multible rib fractures; Clavicle or scapula fracture | 64 | 1.5 | 35.9 | 59 [43, 66] | 81.2 |
| ibip63 | Face fracture other; TBI Concussion | 280 | 1.5 | 84.6 | 43 [23, 57] | 65.7 |
| ibip64 | Thoracic or lumbar spinal cord; Multible rib fractures | 62 | 1.4 | 17.7 | 52 [39, 66] | 61.3 |
| ibip65 | Thoracic or lumbar spine fracture; Lung pneumo or haemothorax; Clavicle or scapula fracture | 14 | 1.4 | 21.4 | 26 [20, 32] | 92.9 |
| ibip66 | Single rib fracture; Clavicle or scapula fracture | 167 | 1.4 | 72.5 | 52 [41, 60] | 79.0 |
| ibip67 | Hand fracture; Radius fracture distal or shaft | 165 | 1.4 | 64.2 | 42 [21, 57] | 67.9 |
| ibip68 | Malar or maxillary fracture; TBI Concussion | 92 | 1.4 | 85.9 | 47 [34, 59] | 64.1 |
| ibip69 | Open wound; TBI Concussion | 90 | 1.2 | 74.4 | 47 [30, 62] | 65.6 |
| ibip70 | TBI Focal brain injury; TBI SDH; TBI SAH; TBI Skull fracture | 80 | – | 43.8 | 56 [32, 67] | 66.2 |
| ibip71 | Thoracic or lumbar spine fracture; Multible rib fractures; Lung pneumo or haemothorax; Clavicle or scapula fracture | 113 | – | 13.3 | 50 [38, 63] | 78.8 |
| ibip72 | Abdominal internal; Lung other; Multible rib fractures; Lung pneumo or haemothorax | 150 | – | 14.0 | 42 [26, 58] | 78.0 |
| ibip73 | Lung pneumo or haemothorax; Single rib fracture | 132 | – | 51.5 | 50 [34, 59] | 78.0 |
| ibip74 | Hand fracture; Radius fracture proximal | 79 | – | 91.1 | 38 [31, 47] | 67.1 |
| ibip75 | Clavicle or scapula fracture; Shoulder girdle joints dislocation | 88 | – | 68.2 | 46 [37, 60] | 87.5 |
| ibip76 | Muscle or joint; Tibia or fibula fracture | 175 | – | 62.3 | 41 [23, 56] | 66.3 |
| ibip77 | Pelvis fracture other; Thoracic or lumbar spine fracture | 245 | – | 28.2 | 50 [33, 64] | 58.4 |

See Appendix A, Table A1 for detailed description of IBIP. n is the final number of individuals sustaining present IBIP. Lift is computed using people with 2–5 injury categories. Lift for IBIP70 – IBIP77 is not generated from the initial computation. 1-IBIP – Percentage who has only present IBIP assigned; TBI – Traumatic Brain Injury; EDH – Epidural Haemorrhage; SDH – Subdural Haemorrhage; SAH – Subarachnoid Haemorrhage; IQR – Inter Quartile Range.

**Table 2**
Demographics of road users sustaining two or more injury categories.

| Road user type | n (%) | Median age [IQR] | Males [%] |
|---|---|---|---|
| All road users | 12,064 | 48 [27, 62] | 65.0 |
| Car occupants | 3666 | 45 [26, 64] | 60.8 |
| Cyclists in collision with another road user | 1409 | 51 [34, 62] | 63.2 |
| Cyclists single | 3357 | 50 [31, 62] | 62.0 |
| Moped riders | 727 | 25 [16, 51] | 84.3 |
| Motorcycle riders | 1547 | 46 [31, 56] | 89.9 |
| Pedestrians in collision with another road user | 1030 | 58 [33, 73] | 45.5 |
| Truck and bus occupants[a] | 328 | 53 [36, 68] | 62.8 |

[a] Excluded in further analysis. IQR – Inter Quartile Range.

recent translation of ICD-10-CM to the AIS08 based on the severity of AIS (Loftis et al., 2016), Glerum and Zonfrillo (2019) found varying levels of agreement between the dictionaries in different body regions, ranging from 86% for the face region to 44% for the head region. Similarly, an Australian study translating the Australian Modification of ICD-10 (ICD-10-AM) to AIS severity suggested there was a low to moderate correlation on individual patient level (Dinh et al., 2020). Collapsing different injury types into the SAID injury categories (Table A1) provides an alternative to existing mapping-tools. Given AIS is generally only recorded with injury data in specialised trauma registries, many previous studies regarding mortality and disability have only been able to use ICD-10 recorded data. Our SAID Injury Taxonomy (Table A1) offers new possibilities for these types of studies designed to compare injury outcome.

Another novel distinction of the approach taken in this work was that the IBIPs were generated from data representing the general road crash trauma population. This not only allowed identification of injury categories including all levels of injury severity, but also meant that injuries occurring across all road user types were used to generate the IBIPs. This is a different approach to that traditionally taken in road safety studies where normal practice is to first divide the data into different road users and then undertake the injury analysis (Santamariña-Rubio et al., 2007).

One methodological issue to note is that among road users with multiple injury categories, 39.8% had injuries that were not significantly associated with other injuries according to our criteria and were not

**Table 3**
IBIPs significantly associated to one road use type (reference = 1). Road user types including five or more individuals are presented.

| IBIP | Individual-Based Injury Pattern (IBIP) | n | Car occupant OR [95%CI] | Cyclist OR [95%CI] | Cyclist_S[a] OR [95%CI] | Moped OR [95%CI] | Motorcycle OR [95%CI] | Pedestrian OR [95%CI] |
|---|---|---|---|---|---|---|---|---|
| ibip30 | Other or unknown; Spine sprain or strain | 44 | 1 | – | 0.29 [0.13, 0.58] | – | – | – |
| ibip40 | Muscle or joint; Spine sprain or strain | 189 | 1 | 0.27 [0.15, 0.45] | 0.08 [0.04, 0.15] | 0.47 [0.25, 0.81] | 0.2 [0.11, 0.34] | 0.2 [0.09, 0.38] |
| ibip41 | Cervical spine fracture; Thoracic or lumbar spine fracture | 182 | 1 | 0.25 [0.12, 0.45] | 0.21 [0.13, 0.33] | 0.47 [0.23, 0.87] | 0.54 [0.35, 0.83] | 0.39 [0.2, 0.68] |
| ibip49 | Lung other; Multible rib fractures | 191 | 1 | 0.28 [0.14, 0.49] | 0.1 [0.05, 0.18] | 0.44 [0.21, 0.81] | 0.6 [0.4, 0.89] | 0.54 [0.32, 0.88] |
| ibip61 | Thoracic or lumbar spine fracture; Multible rib fractures | 218 | 1 | 0.15 [0.07, 0.27] | 0.03 [0.01, 0.07] | 0.17 [0.07, 0.36] | 0.28 [0.18, 0.43] | 0.3 [0.17, 0.49] |
| ibip17 | Radius fracture distal or shaft; Ulna fracture | 996 | 0.2 [0.16, 0.25] | 0.43 [0.34, 0.54] | 1 | 0.58 [0.43, 0.75] | 0.32 [0.25, 0.4] | 0.22 [0.15, 0.3] |
| ibip31 | Radius fracture proximal; Ulna fracture | 186 | 0.28 [0.18, 0.41] | 0.37 [0.21, 0.61] | 1 | 0.44 [0.21, 0.8] | 0.24 [0.13, 0.41] | – |
| ibip63 | Face fracture other; TBI Concussion | 275 | 0.42 [0.31, 0.57] | 0.66 [0.45, 0.96] | 1 | 0.49 [0.27, 0.84] | 0.09 [0.03, 0.18] | 0.58 [0.37, 0.89] |
| ibip36 | Lung other; Multible rib fractures; Clavicle or scapula fracture | 92 | 0.53 [0.32, 0.87] | 0.33 [0.14, 0.68] | 0.15 [0.07, 0.3] | – | 1 | 0.37 [0.15, 0.79] |
| ibip42 | Abdominal internal; Multible rib fractures; Clavicle or scapula fracture | 118 | 0.39 [0.25, 0.6] | 0.15 [0.06, 0.32] | – | 0.24 [0.08, 0.54] | 1 | 0.49 [0.26, 0.86] |
| ibip8 | Acetabulum fracture; Pelvis fracture other | 171 | 0.38 [0.26, 0.58] | 0.26 [0.14, 0.46] | 0.18 [0.11, 0.28] | – | 0.22 [0.12, 0.38] | 1 |
| ibip10 | TBI Focal brain injury; TBI Skull fracture | 100 | 0.29 [0.17, 0.52] | 0.52 [0.27, 0.96] | 0.25 [0.14, 0.44] | – | 0.17 [0.07, 0.37] | 1 |
| ibip35 | Pelvis fracture other; Multible rib fractures; Lung pneumo or haemothorax | 184 | 0.44 [0.31, 0.63] | 0.16 [0.08, 0.3] | 0.04 [0.01, 0.07] | – | 0.32 [0.2, 0.51] | 1 |
| ibip52 | Face fracture other; TBI Skull fracture | 123 | 0.44 [0.27, 0.74] | 0.47 [0.24, 0.87] | 0.33 [0.19, 0.56] | – | 0.19 [0.08, 0.39] | 1 |
| ibip56 | Pelvis fracture other; Multible rib fractures; Clavicle or scapula fracture | 105 | 0.29 [0.18, 0.47] | 0.25 [0.13, 0.48] | 0.07 [0.03, 0.14] | – | 0.22 [0.11, 0.41] | 1 |
| ibip77 | Pelvis fracture other; Thoracic or lumbar spine fracture | 237 | 0.4 [0.29, 0.56] | 0.24 [0.14, 0.38] | 0.02 [0.01, 0.04] | 0.18 [0.08, 0.35] | 0.37 [0.25, 0.55] | 1 |

[a] Cyclist_S: Cyclist Single, not in collision with another road user. TBI – Traumatic Brain Injury.

assigned an IBIP. Further analysis into whether this occurs more often in any specific road user type is needed but beyond the scope of the work presented here. Another issue is whether our decision to only use people with 2 – 5 injury categories (92.2% of the sample) to identify IBIPs in the data mining analysis could influence the proportion of individuals who are assigned a pattern. However, it is a trade-off between the gain in the number of individuals assigned a pattern and the number of patterns identified. We conducted computations using the full sample of multiply injured individuals (n = 12,064) to confirm the maximum number of available patterns. This resulted in 758 patterns with many patterns generated from the same individuals with many injuries, hence the proportion of road users assigned a pattern only increased to 63.6% (n = 7671). However, all but twelve road users with more than five coded injury categories (up to 19 in this study) were assigned a pattern with our presented methodology. Despite the limitation introduced by this trade-off, our method allowed a general set of injury patterns to be generated which can be used in further analyses.

Finally, it is important to note that IBIPs generated within the same body region are dependent on the number of available injury categories in that body region. In the categorisation process described in Appendix A, Step 2, traumatic brain injuries were kept separate in the SAID taxonomy (Table A1). The many different significant associations of TBI injuries suggest that it may be important to analyse their patterns separately rather than on one collapsed TBI group (Maas et al., 2008). Another cluster of injuries that generated numerous patterns of similar injury categories were those in the torso region. However, these injuries originate from four different body regions including the abdomen; thorax; spine; and upper extremities. It is possible to select subset-patterns of these IBIPs but the detail will decline. This is shown with the addition of IBIP71 that resulted in a decrease of IBIP65 to only include 14 people. If a subset-pattern (e.g. IBIP65) is selected as the main pattern it may in itself be very rare. We acknowledge that further research into these IBIPs may suggest that significant subset-patterns could in some cases be a better option. We also acknowledge that the single injury categories or the excluded minor injuries such as "TBI Concussion" and "Spine sprain or strain" may impact long-term outcomes.

The next step of this research is to quantify the unified injury outcome of the IBIPs in terms of mortality and disability. Furthermore, it is our intention to present a tool that can recommend a pathway to analyse road crashes which can contribute to take informed decisions for crash and injury prevention in road traffic.

## 5. Limitations

The limitations are related to the development of injury categories presented in Appendix A which are the input to the association rule mining of IBIPs. The published empirical data used in this process, i.e. from the Validating and Improving Injury Burden Estimates Study (Injury-VIBES) (Gabbe et al., 2016) and the Model Average Regression Coefficient (MARC) values (used in the Trauma Mortality Prediction Model) (Osler et al., 2019), was not available in its original format and therefore computations included some assumptions. However, we consider the computations satisfactory but acknowledge that they might be improved if original data was used (Appendix A).

## 6. Conclusions

An injury taxonomy was developed, named SAID, which offers possibilities for analyses of data recorded in either ICD-10 or AIS08 to be compared. A novel application of a data mining technique was successfully used to identify IBIPs of associated co-occurring injury categories and some clear distinctions in IBIPs between different road user types were demonstrated.

## CRediT authorship contribution statement

All authors were involved in the study concept and design. HF acquired, processed, and analysed the data. HF drafted the manuscript. All authors were involved in interpretation of data and critical revision of the manuscript. All authors read and approved the final manuscript.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors wish to thank the Swedish Transport Agency for providing access to the data. Thanks to Dr. Susan Adams FRACS for expert clinical advice and Dr. Tom Whyte for valuable discussion on brain injuries. Thank you Dr. Nancy Briggs and Dr. Raymond Wong for guidance in data mining and Andrew Kirk for support in R programming. Thanks to Gunilla Collin for expert communication on AIS data input.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aap.2021.106479.

## References


AAAM, 2008. Abbreviated Injury Scale (AIS) 2005 - Update 2008. Association for the Advancement of Automotive Medicine, Barrington, IL.

Agrawal, R., Imielinski, T., Swami, A., 1993. Mining Association in Large Databases. Proc. 1993 ACM SIGMOD Int. Conf. Manag. data - SIGMOD '93 207–216. doi: 10.1145/170036.170072.

Aharonson-Daniel, L., Boyko, V., Ziv, A., Avitzour, M., Peleg, K., 2003. A new approach to the analysis of multiple injuries using data from a national trauma registry. Inj. Prev. 9, 156–162.

Aharonson-Daniel, L., Giveon, A., Peleg, K., 2005. Gaps in injury statistics: multiple injury profiles reveal them and provide a comprehensive account. Inj. Prev. 11 (4), 197–200. https://doi.org/10.1136/ip.2005.008227.

Beck, B., Cameron, P.A., Fitzgerald, M.C., Judson, R.T., Teague, W., Lyons, R.A., Gabbe, B.J., 2017. Road safety: serious injuries remain a major unsolved problem. Med. J. Aust. 207 (6), 244–249. https://doi.org/10.5694/mja17.00015.

Dinh, M.M., Singh, H., Sarrami, P., Levesque, J.-F., 2020. Correlating injury severity scores and major trauma volume using a state-wide in-patient administrative dataset linked to trauma registry data—a retrospective analysis from New South Wales Australia. Injury 51 (1), 109–113. https://doi.org/10.1016/J.INJURY.2019.09.022.

Fagerlind, H., Harvey, L., Candefjord, S., Davidsson, J., Brown, J., 2019. Does injury pattern among major road trauma patients influence prehospital transport decisions regardless of the distance to the nearest trauma centre? – a retrospective study. Scand. J. Trauma. Resusc. Emerg. Med. 27 (18), 1–9. https://doi.org/10.1186/s13049-019-0593-7.

Gabbe, B.J., Lyons, R.A., Fitzgerald, M.C., Judson, R., Richardson, J., Cameron, P.A., 2015. Reduced population burden of road transport-related major trauma after introduction of an inclusive trauma system. Ann. Surg. 261 (3), 565–572. https://doi.org/10.1097/SLA.0000000000000522.

Gabbe, B.J., Lyons, R.A., Simpson, P.M., Rivara, F.P., Ameratunga, S., Polinder, S., Derrett, S., Harrison, J.E., 2016. Disability weights based on patient-reported data from a multinational injury cohort. Bull. World Health Organ. 94 (11), 806–816C. https://doi.org/10.2471/BLT.16.172155.

Gabbe, B.J., Simpson, P.M., Lyons, R.A., Ameratunga, S., Harrison, J.E., Derrett, S., Polinder, S., Davie, G., Rivara, F.P., Reddy, H., 2014. Association between the Number of Injuries Sustained and 12-Month Disability Outcomes: Evidence from the Injury-VIBES Study. PLoS ONE 9 (12), e113467. https://doi.org/10.1371/journal.pone.0113467.

Glerum, K.M., Zonfrillo, M.R., 2019. Validation of an ICD-9-CM and ICD-10-CM map to AIS 2005 update 2008. Inj. Prev. 25 (2), 90–92. https://doi.org/10.1136/injuryprev-2017-042519.

Hahsler, M., 2015. A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules [WWW Document]. URL http://michael.hahsler.net/research/association_rules/measures.html (accessed 1.12.20).

Hahsler, M., Grun, B., Hornik, K., 2005. arules – a computational environment for mining association rules and frequent item sets. J. Stat. Softw. 14 (15), 1–6.

Halpin, J., Greenspan, A.I., Haileyesus, T., Annest, J.L., 2009. The effect of counting principal and secondary injuries on national estimates of motor vehicle-related trauma: a NEISS-AIP special study. Inj. Prev. 15 (5), 328–333. https://doi.org/10.1136/ip.2009.021691.

Howard, C., Linder, A., 2014. Review of Swedish experiences concerning analysis of people injured in traffic accidents. Linköping.

Kotu, V., Deshpande, B., 2019. Association Analysis. In: Data Science. Morgan Kaufmann, pp. 199–220, 10.1016/B978-0-12-814761-0.00006-X.

Liman, S.T., Kuzucu, A., Tastepe, A.I., Ulasan, G.N., Topcu, S., 2003. Chest injury due to blunt trauma. Eur. J. Cardio-Thoracic Surg. 23 (3), 374–378. https://doi.org/10.1016/s1010-7940(02)00813-8.

Loftis, K.L., Price, J.P., Gillich, P.J., Cookman, K.J., Brammer, A.L., St Germain, T., Barnes, J., Graymire, V., Nayduch, D.A., Read-Allsopp, C., Baus, K., Stanley, P.A., Brennan, M., 2016. Development of an expert based ICD-9-CM and ICD-10-CM map to AIS 2005 update 2008. Traffic Inj. Prev. 17 (Suppl 1), 1–5. https://doi.org/10.1080/15389588.2016.1191069.

Maas, A.I., Stocchetti, N., Bullock, R., 2008. Moderate and severe traumatic brain injury in adults. Lancet Neurol. 7 (8), 728–741. https://doi.org/10.1016/S1474-4422(08)70164-9.

Monárrez-Espino, J., Laflamme, L., Berg, H.-Y., 2018. Measuring and assessing risk of quality of life loss following a road traffic injury: a proposed methodology for use of a composite score. Accid. Anal. Prev. 115, 151–159. https://doi.org/10.1016/J.AAP.2018.02.009.

Osler, T., Baker, S., Long, W., 1997. A modification of the Injury Severity Score that both improves accuracy and simplifies scoring. J. Trauma Inj. Infect. Crit. Care 43 (6), 922–926. https://doi.org/10.1017/CBO9781107415324.004.

Osler, T.M., Glance, L.G., Cook, A., Buzas, J.S., Hosmer, D.W., 2019. A trauma mortality prediction model based on the ICD-10-CM lexicon: TMPM-ICD10. J. Trauma Acute Care Surg. 86 (5), 891–895. https://doi.org/10.1097/ta.0000000000002194.

Polinder, S., Haagsma, J., Bos, N., Panneman, M., Wolt, K.K., Brugmans, M., Weijermars, W., van Beeck, E., 2015. Burden of road traffic injuries: Disability-adjusted life years in relation to hospitalization and the maximum abbreviated injury scale. Accid. Anal. Prev. 80, 193–200. https://doi.org/10.1016/j.aap.2015.04.013.

R Core Team, 2018. R: A Language and Environment for Statistical Computing.

RStudio Team, 2016. RStudio: Integrated Development Environment for R.

Santamariña-Rubio, E., Pérez, K., Ricart, I., Arroyo, A., Castellà, J., Borrell, C., 2007. Injury profiles of road traffic deaths. Accid. Anal. Prev. 39 (1), 1–5. https://doi.org/10.1016/J.AAP.2006.06.019.

Talbot, B.S., Gange, C.P., Chaturvedi, A., Klionsky, N., Hobbs, S.K., Chaturvedi, A., 2017. Traumatic rib injury: patterns, imaging pitfalls, complications, and treatment. Radiographics 37 (2), 628–651. https://doi.org/10.1148/rg.2017160100.

Tournier, C., Charnay, P., Tardy, H., Chossegros, L., Carnis, L., Hours, M., 2014. A few seconds to have an accident, a long time to recover: consequences for road accident victims from the ESPARR cohort 2 years after the accident. Accid. Anal. Prev. 72, 422–432. https://doi.org/10.1016/j.aap.2014.07.011.

Vos, T., Abajobir, A.A., Abate, K.H., Abbafati, C., Abbas, K.M., Abd-Allah, F., 2017. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 390 (10100), 1211–1259. https://doi.org/10.1016/S0140-6736(17)32154-2.

WHO, 2016. International Statistical Classification of Diseases and Related Health Problems. 10th Revision [WWW Document]. URL https://icd.who.int/browse10/2016/en (accessed 2.11.19).