



Australian Government

**Australian Institute of
Health and Welfare**

*Better information and statistics
for better health and wellbeing*

Injury severity scaling

A comparison of methods for measurement of injury severity

Geoff Henley, James E Harrison



Flinders
UNIVERSITY

AIHW INJURY TECHNICAL PAPER SERIES NO. 10

ISBN 978 1 74024 887 7



Australian Government

**Australian Institute of
Health and Welfare**

*Better information and statistics
for better health and wellbeing*

INJURY TECHNICAL PAPER SERIES

Number 10

Injury severity scaling

A comparison of methods for measurement of injury severity

**Geoff Henley
James Harrison**

June 2009

Australian Institute of Health and Welfare
Canberra

Cat. no. INJCAT 126

The Australian Institute of Health and Welfare is Australia's national health and welfare statistics and information agency. The Institute's mission is better information and statistics for better health and wellbeing.

© Australian Institute of Health and Welfare 2009

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced without prior written permission from the Australian Institute of Health and Welfare. Requests and enquiries concerning reproduction and rights should be directed to the Head, Media and Communications Unit, Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601.

This publication is part of the Australian Institute of Health and Welfare's Injury Research and Statistics Series. A complete list of the Institute's publications is available from the Institute's website <www.aihw.gov.au>. Electronic copies of publications in this series can be downloaded from the Research Centre for Injury Studies website <www.nisu.flinders.edu.au>

ISSN 1444-778X

ISBN 978 1 74024 887 7

Suggested citation

Australian Institute of Health and Welfare: Henley G & Harrison JE 2009. Injury severity scaling: A comparison of methods for measurement of injury severity. Injury technical paper series no. 10. Cat. no. INJCAT 126. Canberra: AIHW.

Australian Institute of Health and Welfare

Board Chair

Hon. Peter Collins, AM, QC

Director

Penny Allbon

Any enquiries about or comments on this publication should be directed to:

Geoff Henley

Research Centre for Injury Studies

Flinders University

GPO Box 2100

Adelaide SA 5001

Phone: (08) 8201 7602

Email: Geoffrey.Henley@flinders.edu.au

Published by the Australian Institute of Health and Welfare

Proof reading & layout editing by Stacey Avefua

**Please note that there is the potential for minor revisions of data in this report.
Please check the online version at <www.aihw.gov.au> for any amendments.**

Contents

1	Introduction.....	1
2	Methods	3
2.1	Data sources.....	3
2.2	Case definition.....	3
2.3	ICISS.....	3
2.4	Comorbidity index.....	4
2.5	Statistical analysis	5
3	Models including only age as a co-predictor	6
3.1	Multiplicative model	6
3.2	Worst injury model.....	6
3.3	Hybrid models.....	6
3.4	Results.....	7
4	Models including only comorbidity as a co-predictor	10
4.1	Multiplicative model	10
4.2	Worst injury model.....	10
4.3	Results.....	10
5	Models including age and comorbidity as co-predictors	12
5.1	Multiplicative model	12
5.2	Worst injury model.....	12
5.3	Hybrid models.....	12
5.4	Results.....	12
6	Discussion	15
	Appendix A1: Inclusion of same-day cases	17
	Appendix A2: Charlson to ICD-10-AM map	19
	References	20
	List of tables	21
	List of figures	21

1 Introduction

Injury severity measurement is important for meaningful comparison of outcomes of trauma care, and assessment of burden of injury. Trends in hospital separations are often used as a proxy for trends in population incidence. These trends can be misleading due to variations over time or between places in factors unrelated to incidence, such as admission policies and practices. Injury severity measurement is a promising basis for improving measurement of trends in the population incidence of injury, as the incidence of serious injury is less likely to be affected by extraneous factors (Cryer & Jarvis 1999).

Different injury severity criteria are potentially of interest depending on the application. These include threat-to-life (probability of death), cost, impairment, incapacity, impact on quality of life, and disability (Langley & Cryer 2000; Cryer & Langley 2002).

Two methods have been shown to be capable of providing fairly reliable probability of death estimates based on ICD-9-CM diagnosis codes. One method involves translation of ICD-9-CM codes to AIS scores (AAAM 1990) via the proprietary software package ICDMAP-90 (Mackenzie & Sacco 1997). The Abbreviated Injury Scale (AIS) has long been regarded as the injury coding 'industry standard' with respect to its injury-specific descriptive abilities. The other method, known as ICISS (ICD-based Injury Severity Score), involves calculating a Survival Risk Ratio (SRR), i.e. the probability of survival, for each individual injury diagnosis code as the ratio of the number of patients with that injury code who have not died to the total number of patients diagnosed with that code. Thus, a given SRR represents the likelihood that a patient will survive a particular injury. Each patient's ICISS score (survival probability) is then the product of the probabilities of surviving each of their injuries individually. This may be a single SRR, as in the case of a patient with a single injury, or it may be multiple SRRs, as in the case of a patient with multiple injuries (Osler et al. 1996).

More recent evidence suggests ICISS scores derived from ICD-10-AM coding provide a reasonable estimate of injury severity (Stephenson et al. 2003). However, addressing a number of limitations may subsequently improve the overall performance of the ICISS method. These limitations include:

- Inclusion of deaths which only occur in hospital
- Inability to identify individuals who underwent more than one separation within the study period
- Non-independent SRRs
- Exclusion of non-injury diagnoses (i.e. comorbidity)
- Some research suggests that setting ICISS to the same value as the worst injury SRR may be a better approach than multiplying all injury SRRs to obtain ICISS (Kilgo et al. 2003).

Excluding deaths that occur outside hospital detracts from the value of the resulting severity estimates whilst an inability to identify individuals undergoing more than one separation can result in an overestimation of cases in relation to the number of deaths. Both of these limitations can be overcome to some extent by the introduction of data linkage. A system to do this already exists in New Zealand and other systems are in place such as the Western Australian Data Linkage System (WADLS) and the Centre for Health Record Linkage (CHeReL) located in NSW. The introduction of such a system in the rest of Australia would significantly improve the accuracy of estimates for person-based injury incidence as well as providing the ability to distinguish between deaths that occur either in or outside of hospital.

A constraint on this approach is that the ICD-10 injury codes in death records tend to be lower in number and less specific than those in hospital separation records.

The use of non-independent SRRs can lead to an underestimation of survival. This is due to the fact that data from more serious injuries contribute the appearance of severity to the SRRs of minor injuries when they are calculated from patients that have multiple injuries.

The limitation relating to the exclusion of non-injury diagnoses means that the affect of comorbidity is not taken into consideration. Pre-existing conditions such as heart or respiratory disease can significantly impact upon an individual's capacity to survive certain types of injury. Hence the inclusion of non-injury diagnoses into any injury severity scoring system can improve the ability to discriminate between individuals who are likely to die and those that are likely to survive. A recent study assessed models which incorporated comorbidity and suggested that the predictive ability of ICISS would be improved if comorbidity data were included in its calculation (Davie et al. 2008).

We have previously published the SRRs that were produced in a paper in which we and colleagues from New Zealand applied the ICISS method to ICD-10-AM coded data from both countries (Stephenson et al. 2003). These SRRs were based on hospital separations data for the period from 1 July 1999 to 30 June 2001.

We now see reasons to generate a new set of SRRs:

1. With passage of time, a set of SRRs may become a less reliable guide to the probability of death, due to changes on case outcomes.
2. The set of SRRs published previously were calculated using a set of injury separation records that excluded same-day cases, except those that ended with death. Most NISU (National Injury Surveillance Unit) reports use selection criteria that do not omit same-day cases. It is convenient to have a set of SRRs that is based upon our usual selection criteria.

The findings discussed above concerning methods by which SRRs are used to generate case-specific predicted probabilities of survival (i.e. ICISS) also prompts us to test the effects of some variations of method when applied to recent Australian data. The purpose of this report is to demonstrate the effect on the accuracy of estimation of mortality by adjusting for the effect of including non-injury diagnoses (comorbidity).

2 Methods

2.1 Data sources

Australian hospital separations data were used for this project. Data were from the Australian Institute of Health and Welfare (AIHW) National Hospital Morbidity Database (NHMD). Hospital separations were from 1 July 2005 to 30 Jun 2007.

2.2 Case definition

Separations that satisfied all of the following criteria were selected:

A 'Principal Diagnosis' in the ICD-10-AM range S00-T89 (i.e. Chapter XIX Injury, poisoning and certain other consequences of external causes, except for sequelae of injury).

Separation mode was coded to 'Transfer-nursing', 'Transfer-psychiatric', 'Transfer-other health', 'Discharge own risk', 'Statistical discharge-leave', 'Died' or 'Other-usual residence'.

Separations for which separation mode was coded to either 'Transfer-other acute' or 'Statistical discharge-type change' were omitted in order to reduce multiple counting of injury cases (i.e. cases in which there was more than one separation event for the same injury). Although the dataset contained records from both the 4th and 5th editions of ICD-10-AM, no mapping was required since there were no code changes in the range S00-T89 between these two editions.

2.3 ICISS

The ICISS method involves calculating a Survival Risk Ratio (SRR), i.e. the probability of survival, for each individual injury diagnosis code as the ratio of the number of patients with that injury code who have not died to the total number of patients diagnosed with that code. Thus, a given SRR represents the likelihood that a patient will survive a particular injury. SRRs are used as the basis for calculating a predicted probability of survival to discharge for cases in a study set of hospital separation records in which injury diagnoses are coded in the same way as the dataset used to generate SRRs. In the original work on this topic, each patient's ICISS score (survival probability) is the product of the probabilities of surviving each of their injuries individually. Thus, the ICISS may be a single SRR, as in the case of a patient with a single injury, or it may be the product of multiple SRRs, in the case of a patient with multiple injuries (Osler et al. 1996). Age is often included as a predictor variable in many studies, usually as a proxy for comorbidity.

A recent study compared the standard multiple-injury ICISS with a worst-injury ICISS which, rather than being the product of the SRRs, was the smallest SRR among the diagnoses for a patient (Kilgo et al. 2003). Results indicated that worst-injury ICISS, discriminated survival better, fitted better, and explained more variance than the multiple-injury ICISS. However, it should be noted that this study only included patients from a large trauma database and used ICD-9 coding.

2.4 Comorbidity index

The ICISS has been shown to provide good measures of threat to life following injury, when based on trauma service data (Osler et al. 1996; Meredith et al. 2002), or on administrative data on hospital cases (Stephenson et al. 2002). ICISS measures are imperfect, and the extent of imperfection differs between injury types. For example, higher concordances have previously been found for mechanical trauma, and especially for head injury, than for other types of injury. Concordances were lower for cases for which the Principal Diagnosis was complications of care (ICD-10-AM T80–T88).

As originally proposed, age is included in models to derive ICISS from SRRs. Age can be seen as being a proxy for resilience and for comorbidity. In previous work, exclusion of age from models reduced concordance values by about 0.05 (Stephenson et al. 2003). A recent study found that the inclusion into various models of comorbidity, based on the Charlson Index, improved the predictive ability of ICISS (Davie et al. 2008).

Charlson et al. defined 30 clinically important comorbidities, of which 19 had adjusted relative risks of a one-year mortality greater than 1.2 from a proportional hazards model (Charlson et al. 1987). Each of these independent predictors was assigned a weight of 1, 2, 3 or 6 based on the magnitude of its adjusted relative risk. These weights were added to produce a total comorbidity score for each patient. The higher the score, the more severe the burden of comorbidity. The Dartmouth-Manitoba version of the Charlson Index (DM-CI) by Romano et al. was the first adaptation of the Charlson Index to administrative databases (Romano et al. 1993). It includes additional conceptually similar ICD diagnoses that were not explicitly listed in Charlson's original list of 19. This was meant to increase the sensitivity of the DM-CI. D'Arcy Holman and colleagues at the School of Population Health, The University of Western Australia, subsequently developed SPSS syntax code which allowed the comorbidity score for each hospital patient to be calculated. This syntax generated 17 yes/no Charlson comorbidity variables for each patient record using the Dartmouth-Manitoba algorithm for administrative data. The syntax also included four additional variables covering other forms of chest pain and ischaemic heart disease, cardiac dysrhythmias, hypertensive diseases and obesity. The complete set of ICD-10-AM codes included within each Charlson variable is contained within Appendix A2.

The Charlson Index has properties that appear to suit it for use to incorporate comorbidity directly into ICISS models:

- Validated in terms of threat to life
- Defined in terms of ICD codes
- Widely used

This paper presents findings of an investigation to assess the potential value of using a method based on the Charlson Index to take account of comorbidity in ICISS analysis of threat to life due to injury. The method used for this report involved calculating an SRR for each of the ICD-10-AM codes that contribute to any of the 21 Charlson categories. This method was adopted based to some degree on the work of Davie et al. who found that calculating comorbidity SRRs at the ICD-10-AM code level produced the best outcomes (Davie et al. 2008).

2.5 Statistical analysis

ICISS was assessed on its ability to predict severity, in terms of threat to life, using logistic regression modelling. Models used ICISS and age predictor variables and survival as the outcome. Age was coded as a categorical variable at five levels (0-14, 15-24, 25-64, 65-79, and 80+). ICISS was calculated both with and without the addition of selected comorbidity diagnosis codes.

There are two areas in which logistic regression models can be assessed: discrimination and calibration (Harrell et al. 1996). Discrimination refers to the ability of the model to distinguish survivors from non-survivors whilst calibration measures how accurately the model predicts the probability of death of a given set of cases. A model may be highly discriminating, but poorly calibrated or, alternatively, well calibrated but poorly discriminating.

The level of discrimination was assessed using a measure known as concordance (Harrell et al. 1996). Concordance is calculated by comparing the survival probabilities of all possible pairs of survivors and nonsurvivors. A score of 1 is assigned if the survivor has a greater probability of survival; 0.5, where the survival probabilities are equal; or 0, where the nonsurvivor has a greater probability of survival. The concordance is the sum of these values divided by the total number of survivor/nonsurvivor pairs. A value of 0.5 indicates no predictive discrimination, whereas a value of 1.0 indicates perfect separation of patients with different outcomes. The concordance has been shown to equal the area under the receiver operating characteristic (ROC) curve.

The relative calibration of the scores was compared using calibration curves (Rowan et al. 1993) and the Hosmer-Lemeshow (H-L) statistic (Hosmer & Lemeshow 1989). Calibration curves are a plot of observed against estimated mortality, with cases grouped by estimated mortality. If the model is perfectly calibrated, observed and estimated mortality will be equal and all points on the calibration curve will lie on a 45-degree line. Deviations from this line show when the model is over- or under-predicting mortality.

The H-L statistic is a summary measure of the calibration of the model. It is also based on comparing the observed and estimated mortality for patients grouped by estimated mortality. The resulting statistic follows a χ^2 distribution, with the degrees of freedom equal to two less than the number of groups. The smaller the H-L statistic, the better the fit, with a perfectly calibrated model having a value of zero. The R^2 is a descriptive goodness-of-fit measure between 0 and 1 that describes the proportion of variance explained by the model. Higher values are better.

Calculation of SRR values was undertaken in SPSS version 14.0. Concordance, H-L statistic, and R^2 calculations were conducted using Stata/SE Version 9.

3 Models including only age as a co-predictor

Three types of models were assessed using only age, as a categorical variable, (refer to Section 2.5) as a co-predictor within a logistic regression model. These models are described below.

3.1 Multiplicative model

This model incorporates the original method of calculating each patient's ICISS by multiplying all of the injury SRRs contained in a patient's record to obtain the final score. This method is detailed in the report *Diagnosis-based injury severity scaling: A method using Australian and New Zealand data coded to ICD-10-AM* (Stephenson et al. 2003). This previous report excluded records where the patient was discharged on the same-day as they were admitted, unless they died in hospital. This approach may introduce some bias into the model in that it excludes many cases where a person is admitted to hospital with only minor injuries. Consequently, models that excluded same-day cases appear more likely to overestimate mortality than models that did not include these cases.

For this reason, the current report includes multiplicative models both with and without same-day cases.

3.2 Worst injury model

For early studies investigating the ICISS approach, severity scores were calculated using all of the patient's injuries. The degree to which multiple injuries contribute to outcome prediction is unclear. Meredith et al. found that the MAXAIS (the largest Abbreviated Injury Scale (AIS) severity in a patient's set of injuries) had better discrimination and was better calibrated than its multiple injury competitors ISS and NISS (Meredith et al. 2002). Kilgo et al. adopted a similar approach using both AIS and ICISS-based measures and concluded that a patient's worst injury discriminates survival better, fits better, and explains more variance than currently used multiple injury scores (Kilgo et al. 2003). However, their analysis was based on data from a trauma registry and excluded injuries related to foreign bodies, burns and traumatic complications. The model used in this report calculated each patient's ICISS by setting it to the same value as the lowest injury SRR value contained within a patient's record

3.3 Hybrid models

These models incorporate features of both the multiplicative and 'worst injury' models. We were tempted to try this approach by our observation of the lack of difference between the multiplicative and 'worst injury' models when applied to Australian data. This is discussed further below.

3.3.1 Method 1

The method involved in creating this model is outlined below:

1. Create variable containing ICD injury code with the lowest injury SRR value.
2. If code starts with 'S0' (i.e. head region), then calculate ICISS by multiplying SRRs corresponding to ICD codes beginning with 'S0'.
3. Repeat the above process for codes starting with 'S1' (neck region), 'S2' (thorax region) & 'S3' (abdomen, lower back, lumbar spine and pelvis regions).
4. If code starts with range 'S4'-'S6' (i.e. upper limb region), then calculate ICISS by multiplying SRRs corresponding to ICD codes beginning with 'S4'-'S6'.
5. Repeat the above process for codes starting with range 'S7'-'S9' (i.e. lower limb region).
6. If code starts with 'T', then set ICISS to same as worst injury ICISS value.

3.3.2 Method 2

The method involved in creating this model is outlined below:

1. From the dataset containing hospital separations for the period from 1 July 2005 to 30 June 2007 create a flag variable for each injured body region.
2. Seven flag variables were created which included the following body regions: head; neck; thorax; abdomen, lower back, lumbar spine and pelvis; upper limb; lower limb; and unspecified body region.
3. Select cases where the patient only sustained an injury to a single body region.
4. Use this set of cases to derive a new set of injury SRRs.
5. Apply this set of SRRs to the dataset as outlined in Section 2 of this report to calculate ICISS for each patient.

3.4 Results

Concordance, H-L statistic and R^2

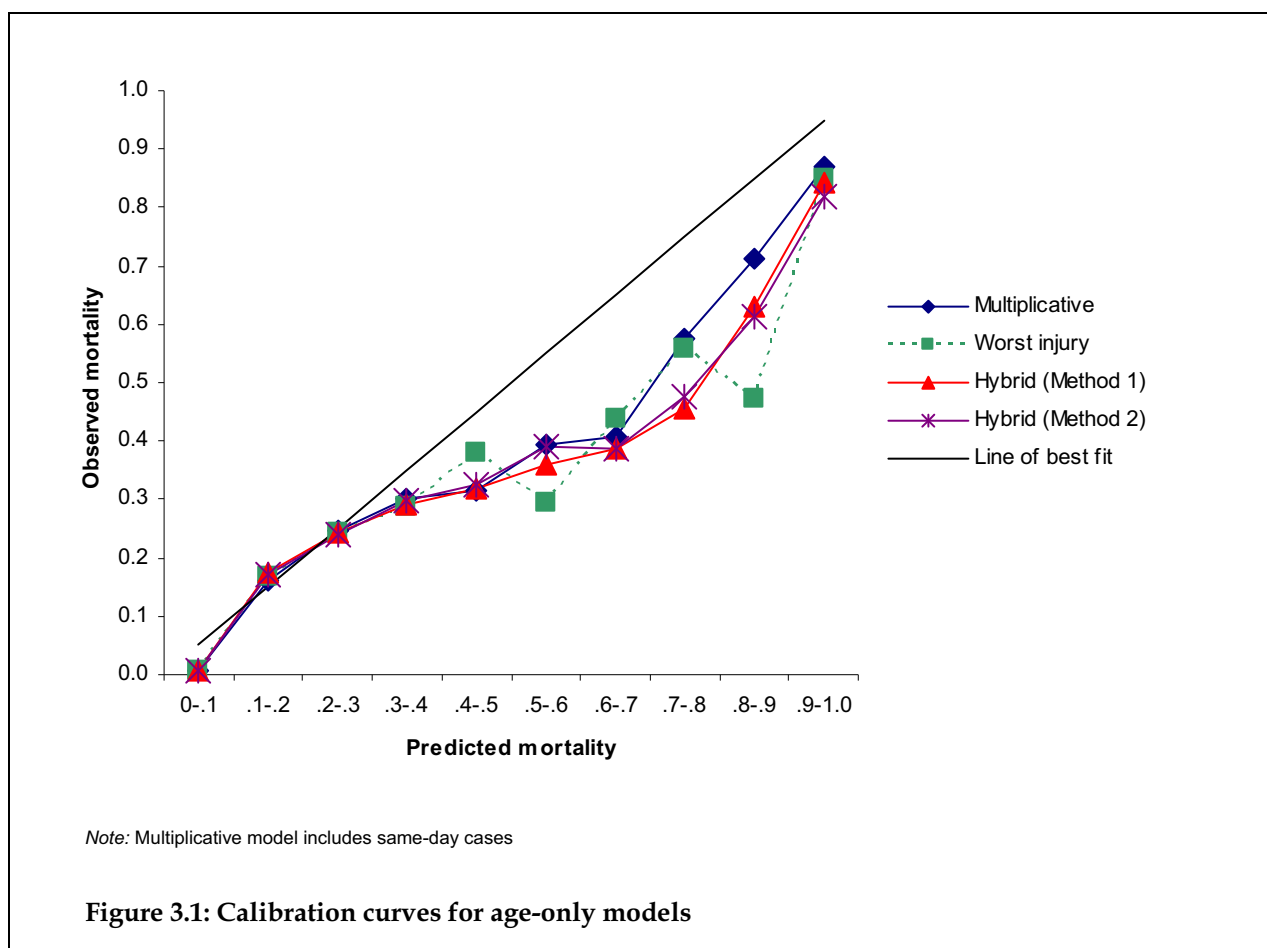
The concordance, Hosmer-Lemeshow (H-L) statistic and R^2 were similar for all five models (Table 3.1). The 'worst injury' model had the highest concordance (0.9262), while the hybrid model (Method 1) had the lowest (0.9218). The H-L statistic was almost identical for the multiplicative and 'worst injury' models with hybrid model (Method 2) having the highest value at 470. The R^2 value suggested that all the models explained around 31%–32% of the variance.

Table 3.1: Concordance and Hosmer-Lemeshow (H-L) statistic and R² for age-only models

Model	Concordance	H-L	R ²
Multiplicative (excluding same-day cases)	0.9237	377	0.3212
Multiplicative (including same-day cases)	0.9232	367	0.3204
Worst injury	0.9262	366	0.3211
Hybrid (Method 1)	0.9218	404	0.3107
Hybrid (Method 2)	0.9223	470	0.3089

Calibration curves

The calibration curves were similar for all models (Figure 3.1). Calibration was generally better at lower mortality, and was particularly good below 30% observed mortality. All models overestimated mortality above this level of observed mortality.



When comparing the multiplicative model using SRRs derived from a dataset including same-day cases with the multiplicative model using SRRs derived from a dataset excluding same-day cases, the resultant calibration curves were almost identical.

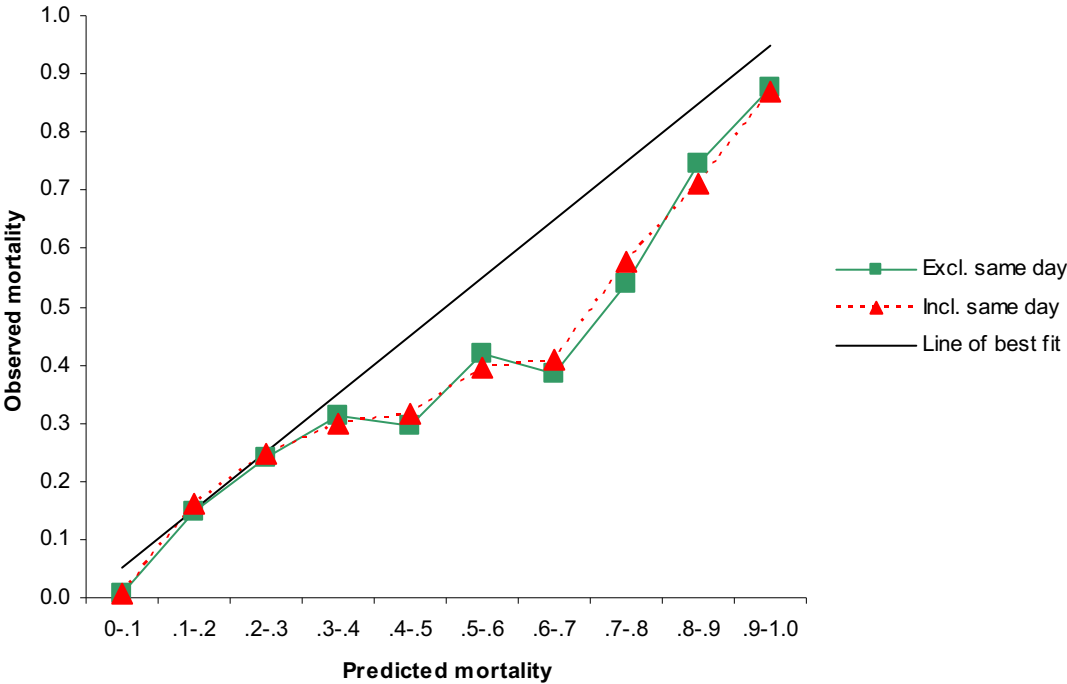


Figure 3.2: Calibration curves for multiplicative models which both excluded and included same-day cases

4 Models including only comorbidity as a co-predictor

Two models were assessed using only comorbidity as a co-predictor within a logistic regression model (i.e. age was excluded from the models). These models are described below.

4.1 Multiplicative model

As outlined in Section 3.1, for the age-only model, only the SRRs corresponding to ICD-10-AM injury codes (i.e. S00-T89) were used in the calculation of ICISS. As outlined in Section 2.4 of this paper, SRRs were also calculated for each of the ICD-10-AM codes that contribute to any of the 21 Charlson categories. Hence, when incorporating comorbidity into the model, the ICISS was calculated by multiplying both sets of SRRs (i.e. injury SRRs and comorbidity SRRs) to obtain the final ICISS value. Consequently, any record which contained one or more of the ICD-10-AM codes belonging to any of the 21 Charlson categories, will have a lower ICISS score than previously, when comorbidity was not included in the model (provided that at least one of the comorbidity codes has a SRR of less than 1).

4.2 Worst injury model

As outlined in Section 3.2, the age-only model calculated each patient's ICISS by setting it to the same value as the lowest injury SRR value contained within a patient's record. For this model, ICISS was calculated by multiplying the lowest SRR corresponding to an injury code, with the product of all the comorbidity SRRs.

4.3 Results

Concordance, H-L statistic and R^2

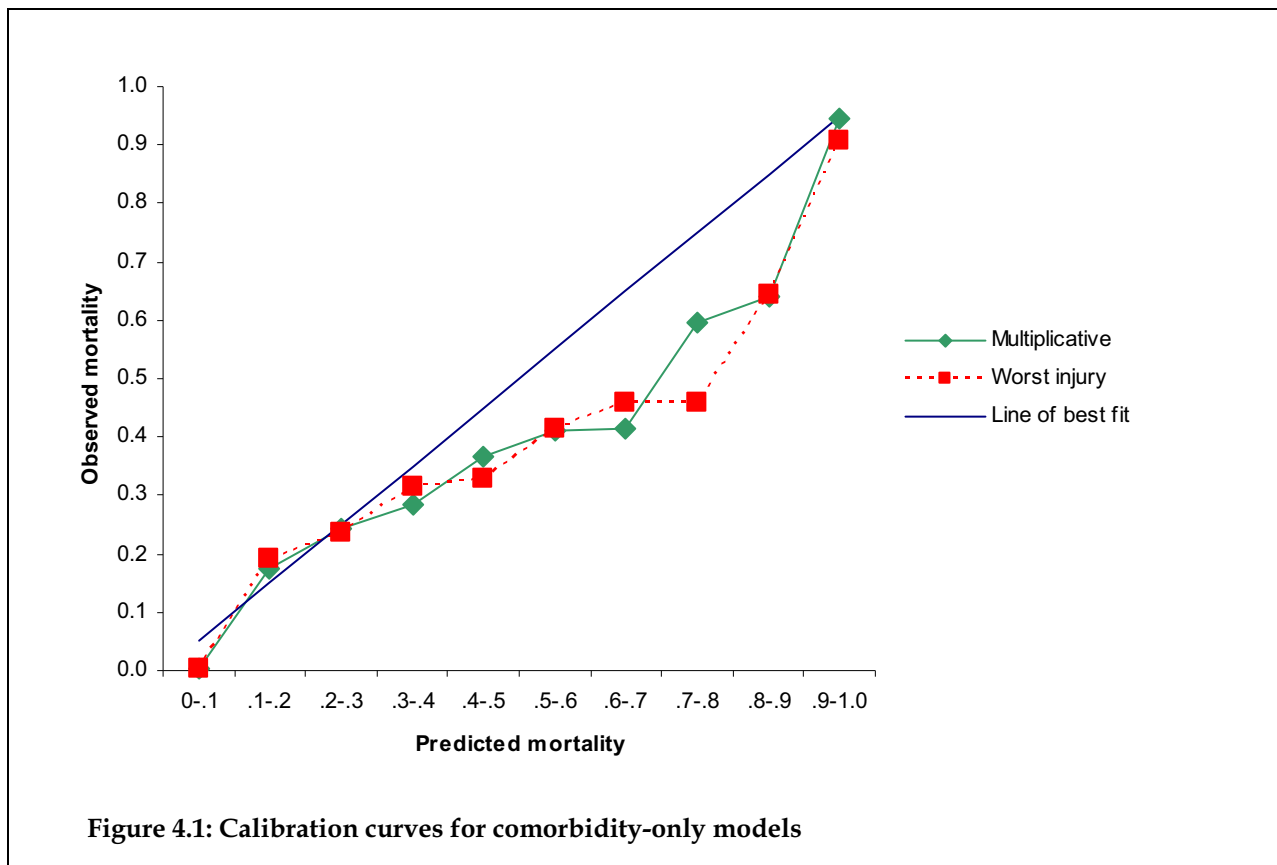
The concordance was similar for both the multiplicative and 'worst injury' models (Table 4.1). Concordance for both models was higher than for the corresponding age-only models (i.e. 0.9646 vs 0.9232 for multiplicative model; 0.9654 vs 0.9262 for 'worst injury' model). Interestingly, the H-L statistic was around three times higher for the comorbidity only models when compared to the age-only models. The R^2 value suggested that both the models explained around 46% of the variance, higher than the 31%-32% for the models with only age as a predictor.

Table 4.1: Concordance and Hosmer-Lemeshow (H-L) statistic and R² for comorbidity-only models

Model	Concordance	H-L	R ²
Multiplicative	0.9646	1100	0.4640
Worst injury	0.9654	1229	0.4570

Calibration curves

The calibration curves were similar for both models (Figure 4.1). As with the age-only models, calibration was generally better at lower mortality, and was particularly good below 30% observed mortality. Both models overestimated mortality above this level of observed mortality.



5 Models including age and comorbidity as co-predictors

Three types of models were assessed using both age and comorbidity as co-predictors within a logistic regression model. These models are described below.

5.1 Multiplicative model

This model is the same as that described in Section 4.1, other than when performing logistic regression, the categorical age variable was included in the model.

5.2 Worst injury model

This model is the same as that described in Section 4.2, other than when performing logistic regression, the categorical age variable was included in the model.

5.3 Hybrid models

The age-only hybrid models incorporating two different methodologies are described in Sections 3.3.1 and 3.3.2. The same methodologies were employed for both models assessed in this section, except that ICISS was calculated by multiplying the previously derived ICISS (i.e. calculated using only injury SRRs) by the product of all the comorbidity SRRs to obtain the final ICISS.

5.4 Results

Concordance, H-L statistic and R^2

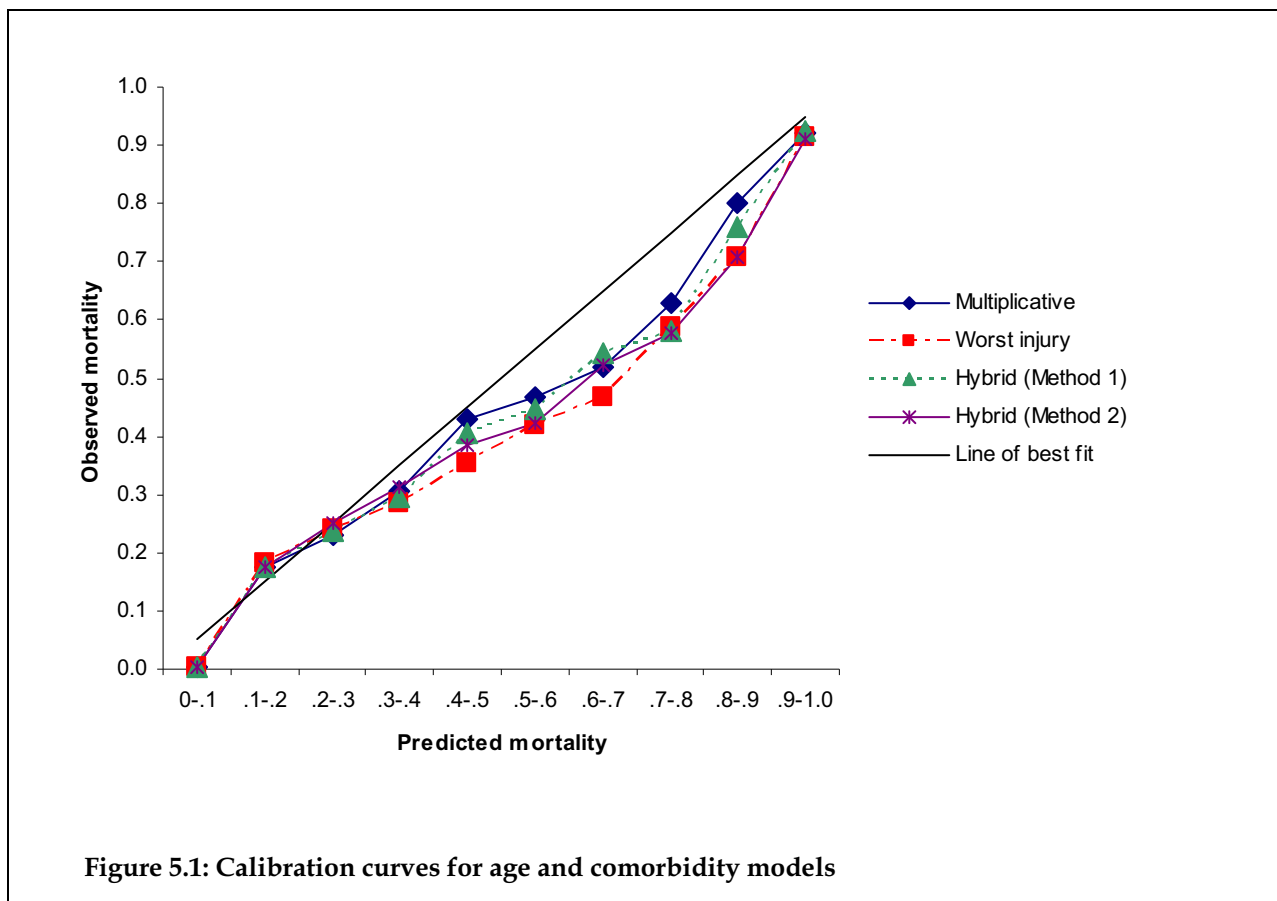
The concordance was similar for all models ranging from 0.9649 from the hybrid (Method 2) model to 0.9675 for the multiplicative model. These values were higher than for the corresponding age-only models where the concordance ranged from 0.9218 to 0.9262. The Hosmer-Lemeshow (H-L) statistic was similar to that observed for the corresponding age-only models, except for the 'worst injury' model where the H-L statistic rose from 366 to 565. The R^2 value suggested that these models explained around 49%–51% of the variance, markedly higher than the 31%–32% for the models with only age as a predictor.

Table 5.1: Concordance, Hosmer-Lemeshow (H-L) statistic and R² for age and comorbidity models

Model	Concordance	H-L	R ²
Multiplicative	0.9675	354	0.5097
Worst injury	0.9667	565	0.4862
Hybrid (Method 1)	0.9668	393	0.5022
Hybrid (Method 2)	0.9649	495	0.4909

Calibration curves

The calibration curves were similar for all models (Figure 5.1). Calibration was generally better at lower mortality, and was particularly good below 40% observed mortality. All models overestimated mortality above this level of observed mortality.



When comparing the original age-only multiplicative model with the age and comorbidity multiplicative model, the calibration curve for the latter was a markedly better fit, particularly above observed mortality of 30% (Figure 5.2).

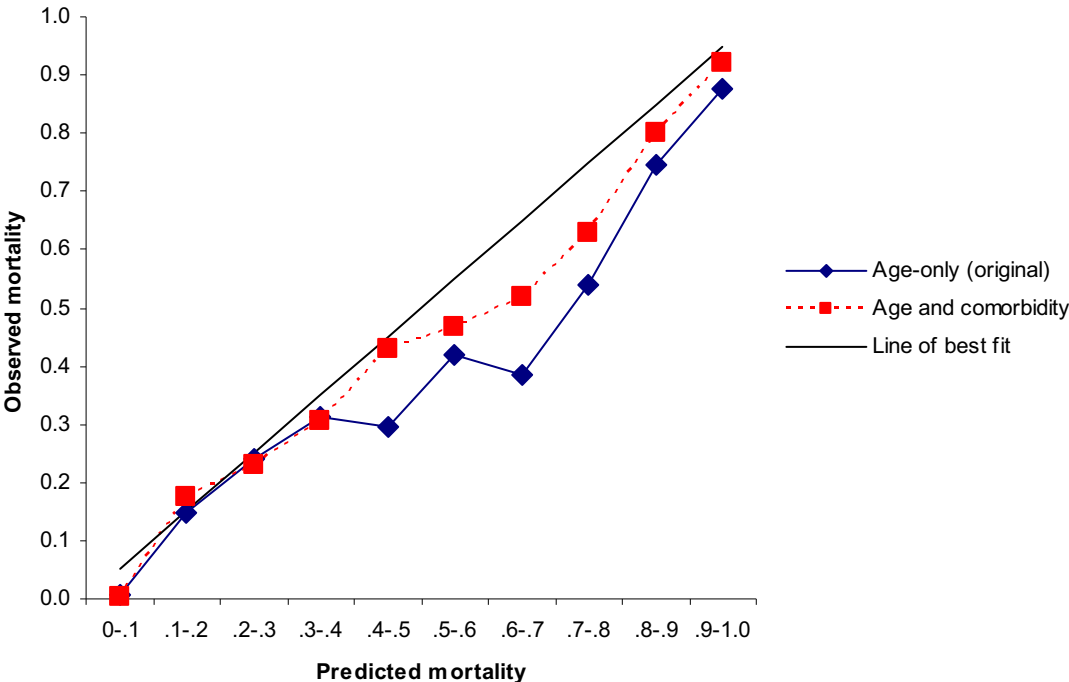


Figure 5.2: Calibration curves for original age-only model and age and comorbidity model

6 Discussion

This report investigated the predictive ability of ICISS by comparing different methods of calculating ICISS. The main methods investigated included the multiplicative and 'worst injury' approaches. Models were also investigated both with and without age and comorbidity as co-predictors.

Age-only models

Comparison of models using the multiplicative and 'worst injury' approaches produced very similar results in terms of both discrimination and calibration. This suggests that there is no distinguishable difference between the two approaches when comparing their ability to predict death, when applied to Australian data. The advantages of the 'worst injury' approach are that it is computationally simpler and that it can be used for datasets where only a relatively small number of injury diagnoses are provided (assuming that the most serious injuries are coded early in the record).

When looking at the multiplicative model, the exclusion of same-day cases did not appear to have any significant effect on the predictive ability on the model. It should be noted that excluding same-day cases will result in deriving SRRs that are either the same as or lower than the equivalent SRRs derived when same-day cases are included. Despite this, both models including and excluding same-day cases appear to overestimate mortality to the same extent.

The adoption of a hybrid approach, utilising two different methodologies, did not improve the predictive ability of the model. This type of approach may require further investigation.

Comorbidity-only models

The incorporation of comorbidity, rather than age, into the models appeared to improve both the multiplicative and 'worst injury' models in terms of discrimination, while the calibration, measured in terms of the H-L statistic, appeared to be worse. However, despite the markedly higher values for the H-L statistic, the calibration curves appear to be similar when comparing the comorbidity-only models with the age-only models. This suggests that the H-L statistic should be treated with caution, and as indicated in previous studies, may be highly sensitive when used with large datasets (Sacco et al. 1999).

Age and comorbidity models

Including both age and comorbidity into both the multiplicative and 'worst injury' models improved model performance in terms of discrimination. However, there was no significant difference in terms of calibration, when measured using the H-L statistic, for the multiplicative model, while for the 'worst injury' model, calibration was slightly worse. However, in contrast to the outcomes suggested by the H-L statistic, calibration curves indicated a better fit for models which included both age and comorbidity. These findings add support to the work of Davie et al. who concluded that the inclusion of comorbidity improves the predictive ability of ICISS (Davie et al. 2008).

From these results, it appears that although age correlates strongly with comorbidity, it does not account for the presence of certain comorbidities that can occur in younger age groups (e.g. obesity, certain types of cancer, paraplegia, and diabetes).

Conclusions

Both the multiplicative and 'worst injury' approaches provide similar outcomes in terms of the predictive ability of ICISS, with the multiplicative model performing slightly better. Hybrid approaches did not appear to improve the predictive ability of any of the models. The inclusion of comorbidity appears to markedly improve model performance. Future possible enhancements incorporate the inclusion of cases where death occurred outside of hospital, and the ability to better identify individuals who underwent more than one hospital separation for a single injury event. Both of these enhancements require further development of data linkage systems. The investigation of modified hybrid approaches may also prove beneficial.

Appendix A1: Inclusion of same-day cases

Section 3.1 refers to the inclusion of same-day cases for the models analysed in this report. Same-day cases are defined as cases in which the patient was discharged on the same-day as admission, unless discharge was due to death. It was felt that excluding same-day cases may introduce some bias into the model in that it excludes many cases where a person is admitted to hospital with only minor injuries. Consequently, it was thought that models that excluded same-day cases were more likely to overestimate mortality than models that did not include these cases.

Comparison of models both including and excluding same-day cases using the multiplicative approach yielded very similar results both in terms of discrimination and calibration (refer to Table 3.1). Hence, it appears that excluding same-day cases does not increase a model's tendency to overestimate mortality. Before discussing the reasons as to why this might be the case, it should be noted that SRRs calculated for each ICD-10-AM code using a dataset which includes same-day cases, will always be equal to or higher than SRRs calculated using an equivalent dataset in which the same-day cases have been removed. Given this fact, it also applies that the ICISS calculated for each individual record using SRRs derived from a dataset including same-day cases will always be equal to or higher than the ICISS calculated using SRRs derived from a dataset excluding same-day cases. To understand why this is the case, it is necessary to look at how SRRs are calculated. The formula is shown below:

$$\text{SRR} = 1 - (n/d)$$

n = the number of records which contain a particular ICD-10-AM code in the dataset where the patient died in hospital

d = the total number of occurrences of that code in the dataset.

The value of n will be the same for both datasets (i.e. including and excluding same-day cases). However, by virtue of the fact that the dataset which includes same-day cases has more records than the dataset that does not, the value of d for the former will always be equal or more than the value of d for the latter. Hence, if d is higher, then (n/d) must be lower and $1 - (n/d)$ must be higher.

With these facts in mind, it is now possible to provide explanations as to why excluding same-day cases does not increase a model's tendency to overestimate mortality:

- ICD-10-AM diagnosis codes for which there is a large difference between the SRR values derived when using datasets which include same-day cases and datasets that exclude them, mostly occur infrequently. Hence, while such differences occur (see ICD-10-AM code characteristics below), they will normally only have a minimal effect in terms of the overall values of ICISS for hospitalised injuries.
- A correlation exists between diagnosis codes which are more likely to be associated with death and the likelihood of being kept in hospital for more than one night. Consequently, codes associated with high threat to life are likely to occur infrequently among same-day cases and SRRs for the same diagnosis codes derived using the two different datasets (i.e. including and excluding same-day cases) are likely to be similar.

- Diagnosis codes which occur frequently in a dataset which excludes same-day cases are generally likely to have a SRR close to 1. Hence, even if these codes occur frequently within the same-day cases, the inclusion of these cases into the dataset will only result in a relatively small increase in the derived value of the SRR for a given code.

It should also be remembered records that involve discharge on the same day as admission, but where the patient is discharged to another acute hospital or involving a statistical-discharge type change, were not included in either of the datasets in order to reduce multiple counting of injury cases (refer to Section 2.2).

ICD-10-AM code characteristics

The characteristics of some ICD-10-AM diagnosis codes can vary depending on whether or not they are included in same-day records. For example, when the code *S09.7 Multiple injuries of head* appears in a same-day record it is nearly always the code which corresponds to the lowest SRR. This is also the case where this code appears in non same-day records if the patient survived. However, where the code appears in a non same-day record and the patient died in hospital, this code only corresponds to the lowest SRR approximately 50% of the time. This highlights the dual nature of this code in that it can be used to code for severe head injuries (i.e. intracranial bleeds and a loss of consciousness of lengthy duration) as well as minor head injuries (i.e. minor cuts and abrasions). A similar pattern is also seen for the code *T17.8 Foreign body in other and multiple sites of respiratory tract*. As with *S09.7*, this code can be used for severe cases (i.e. where airway is completely blocked possibly leading to oxygen deprivation and brain damage, as well as non-severe cases where the patient's breathing is not affected).

In some instances, the SRRs of diagnosis codes corresponding to injuries which would be considered to be non-life threatening can be significantly affected (i.e. reduced) by the presence of other diagnosis codes corresponding to more serious injuries within the same hospital record. For example, the SRR for the diagnosis code *S52.9 Fracture of forearm, part unspecified* when derived using hospital records including same-day cases for the period from July 2005 to June 2007 has a value of 0.9020, whereas the SRR derived for this code when derived using hospital records excluding same-day cases has a value of 0.7826. This is because within the dataset excluding same-day cases, this code is more likely to be associated with other diagnosis codes corresponding to other serious head injuries such as intracranial injury.

Appendix A2: Charlson to ICD-10-AM map

Table A2.1: Charlson diagnostic category to ICD-10-AM code map

Diagnostic category	Revision 1 ICD-10-AM codes
Myocardial infarction	I21.xx, I25.2
Congestive heart failure	I11.0, I13.0, I13.2, I42.xx, I43.xx, I50.xx, I51.7
Peripheral vascular disease	I70.xx, I71.xx, I72.xx, I73.1, I73.8, I73.9
Cerebrovascular disease	G45.0, G45.1, G45.2, G45.8, G45.9, H34.0, I60.xx, I62.xx, I63.xx, I64, I65.xx, I66.xx, R47.0
Dementia	F00.xx, F01.xx, F02.xx, F03, G30.xx, G31.0, G31.1
Chronic pulmonary disease	I26.0, I27.0, I27.8, I27.9, J41.xx, J42, J43.xx, J44.xx, J45.xx, J46, J47
Rheumatologic disease	M05.xx, M06.0x, M06.1x, M06.3x, M06.4x, M06.9x, M08.0x, M08.2x, M08.3, M08.4x, M32.xx, M33.xx, M34.xx, M35.xx
Peptic ulcer disease	K25.x, K26.x, K27.x, K28.x
Mild liver disease	K70.2, K70.3, K72.1, K73.0, K73.1, K73.8, K73.9, K74.x, K76.0
Diabetes (mild to moderate)	E10.0x, E10.1x, E10.60, E10.9x, E11.0x, E11.1x, E11.60, E11.9x, E12.0x, E12.1x, E12.60, E12.9x, E13.0x, E13.1x, E13.60, E13.9x, E14.0x, E14.1x, E14.60, E14.9x
Diabetes with chronic complications	E10.2x–E10.8x, E11.2x–E11.8x, E12.2x–E12.8x, E13.2x–E13.7x, E14.2x–E14.7x.
Hemiplegia and paraplegia	G81.x, G82.00–G82.02, G82.10–G82.12, G82.20–G82.22, G82.30–G82.32, G82.40–G82.42, G82.50–G82.52, G83.0
Renal disease	N18.xx, N19, Z49.1, Z49.2, Z94.0, Z99.2
Any malignancy, including lymphoma and leukaemia	C00.x–C06.x, C07, C08.x–C11.x, C12, C13.x–C18.x, C19, C20, C21.x, C22.x, C23, C24.x–C26.x, C30.x–C32.x, C33, C34.x, C37, C38.x–C41.x, C43.x–C51.x, C52, C53.x, C54.x, C55, C56, C57.x, C58, C60.x, C61, C62.x, C63.x, C64, C65, C66, C67.x–C72.x, C73, C74.x–C76.x, C81.x–C85.x, C88.xx, C90.xx–C95.xx, C96.x, D89.0
Moderate or severe liver disease	I98.xx, I85.x, K72.x, K75.0, K75.1, K76.6, K76.7
Metastatic solid tumor	C77.x, C78.x, C79.xx
AIDS	B20, B21, B22, B23.x, B24
Other ischaemic heart disease	I20.x, I24.x, I25.x, R07.x
Cardiac dysrhythmias	I46.x, I47.x, I48, I49.x
Hypertensive disease	I10, I11.x–I13.x, I15.x
Obesity	E66.x

Source: School of Population Health: University of Western Australia

References

- AAAM (Association for the Advancement of Automotive Medicine) 1990. The abbreviated injury scale. 1990 revision. USA: AAAM.
- Charlson M, Pompei P, Ales K & Mackenzie C 1987. A new method of classifying prognostic comorbidity in longitudinal studies. *Journal of Chronic Diseases*. 40:373–83.
- Cryer C, Langley JD, Stephenson SCR, Jarvis SN and Edwards P 2002. Measure for measure: The quest for valid indicators of non-fatal injury incidence. *Public Health*. 116 (257–262).
- Cryer PC, Jarvis SN, Edwards P & Langley JD 1999. How can we reliably measure the occurrence of non-fatal injury? *International Journal of Consumer and Product Safety*. 6 (4):183–91.
- Davie G, Cryer C & Langley JD 2008. Improving the predictive ability of the ICD-based Injury Severity Score. *Injury Prevention*. 14:250–5.
- Harrell FE Jnr, Lee KL & Mark DB 1996. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*. 15:361–87.
- Hosmer D & Lemeshow S 1989. *Applied Logistic Regression*. New York: John Wiley and Sons, Inc.
- Kilgo PD, Osler TM & Meredith JW 2003. The worst injury predicts mortality outcome the best: rethinking the role of multiple injuries in trauma outcome scoring. *Journal of Trauma: Injury, Infection and Critical Care*. 55 (4):599–607.
- Langley JD & Cryer C 2000. Indicators for injury surveillance. *Australasian Epidemiologist*. 7 (1):5–9.
- Mackenzie EJ & Sacco WJ 1997. *ICDMAP-90 Software : user's guide*. T.J.H. University. USA, The John Hopkins University and Tri-Analytics. 1–25.
- Meredith JW, Evans G, Kilgo PD, Mackenzie EJ, Osler T, McGwin G, Cohn S, Esposito T, Gennarelli T, Hawkins M, Lucas C, Mock C, Rotondo M, & Rue L 2002. A comparison of the abilities of nine scoring algorithms in predicting mortality. *Journal of Trauma: Injury, Infection and Critical Care*. 53(4):621–9.
- Osler T, Rutledge R, Deis J & Bedrick E 1996. ICISS: An International Classification of Disease-9 based Injury Severity Score. *Journal of Trauma: Injury, Infection and Critical Care*. 41 (3):380–8.
- Romano R, Roos L & Jollis J 1993. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *Journal of Clinical Epidemiology*. 46:1075–9.
- Rowan KM, Kerr JH, McPherson K, Short A & Vessey MP 1993. Intensive Care Society's APACHE II study in Britain and Ireland – II: Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *BMJ*. 307:977–81.
- Sacco WJ, Mackenzie EJ, Champion HR, Davis EG & Buckman RF 1999. Comparison of alternative methods for assessing injury severity based on anatomic descriptors. *Journal of Trauma: Injury, Infection and Critical Care*. 47 (3):441-6; discussion 6–7.
- Stephenson C, Henley G, Harrison J & Langley JD 2003. *Diagnosis-based injury severity scaling: A method using Australian and New Zealand data coded to ICD-10-AM*. Cat. no. INJCAT 59. Adelaide: AIHW.
- Stephenson SCR, Langley JD & Civil ID 2002. Comparing measures of injury severity for use with large databases. *Journal of Trauma: Injury, Infection and Critical Care*. 53 (2):326–32.

List of tables

Table 3.1:	Concordance and Hosmer-Lemeshow (H-L) statistic and R^2 for age-only models	8
Table 4.1:	Concordance and Hosmer-Lemeshow (H-L) statistic and R^2 for comorbidity-only models.....	11
Table 5.1:	Concordance, Hosmer-Lemeshow (H-L) statistic and R^2 for age and comorbidity models.....	13
Table A2.1:	Charlson diagnostic category to ICD-10-AM code map.....	19

List of figures

Figure 3.1:	Calibration curves for age-only models.....	8
Figure 3.2:	Calibration curves for multiplicative models which both excluded and included same-day cases.....	9
Figure 4.1:	Calibration curves for comorbidity-only models	11
Figure 5.1:	Calibration curves for age and comorbidity models.....	13
Figure 5.2:	Calibration curves for original age-only model and age and comorbidity model.....	14