Comparison of revised Functional Capacity Index scores with Abbreviated Injury Scale 2008 scores in predicting 12-month severe trauma outcomes

Cameron S. Palmer, GradDip(ClinEpi), BOrth(Hons) [1,2]

Peter A. Cameron, MD, FACEM [1,3]

Belinda J. Gabbe, MBiostat, PhD [1,4]

[1] Department of Epidemiology & Preventive Medicine, Monash University, Melbourne, Australia

[2] Trauma Service, Royal Children's Hospital Melbourne, Parkville, Australia

[3] Emergency & Trauma Centre, The Alfred Hospital, Melbourne, Australia

[4] Health Data Research UK, Swansea University Medical School, Swansea University

Corresponding author:

Mr Cameron S. Palmer

Trauma Service, Royal Children's Hospital Melbourne

50 Flemington Road, Parkville, Victoria 3052, Australia

Telephone: +61 3 9345 4806

E-mail: cameron.palmer@rch.org.au

Keywords:

Functional Capacity Index

Abbreviated Injury Scale

Functional outcomes

12-month outcomes

Trauma scoring

Prediction models

Major trauma

Trauma registry

Word count: 3000 words

KEY MESSAGES

What is already known

- Recovery trajectories following serious injury vary widely, and existing injury severity measures based on Abbreviated Injury Scale (AIS) severity weightings account for only a small proportion of outcome variability.
- The revised predictive Functional Capacity Index (FCI) was developed to predict 12-month outcomes using the AIS code structure, but has not been thoroughly assessed and no summary scores for multiply-injured patients are available.

What this study adds

- Overall anatomical injury as measured by AIS-based or FCI-based scores or a simple injury count all contribute significantly, but only slightly to the prediction of a variety of 12-month outcomes physical and psychosocial outcomes including return to work.
- FCI-based scores do not consistently or substantially improve outcome predictions compared to other injury scores; as such, the FCI is unlikely to be fit for its intended purpose as a global functional outcome prediction tool. Prediction models may require injury scores which are specific to the outcome being assessed.

ABSTRACT

Introduction

Anatomical injury as measured by the Abbreviated Injury Scale (AIS) often accounts for only a small proportion of variability in outcomes after injury. The predictive Functional Capacity Index (FCI) appended to the 2008 AIS claims to provide a widely-available method of predicting 12-month function following injury.

Objectives

To determine the extent to which AIS-based and FCI-based scoring is able to add to a simple predictive model of 12-month function following severe injury.

Methods

Adult trauma patients were drawn from the population-based Victorian State Trauma Registry (VSTR). Major trauma, and severely injured orthopaedic trauma patients were followed up via telephone interview including Glasgow Outcome Scale - Extended (GOS-E), the EQ-5D-3L and return to work status. A battery of AIS-based and FCI-based scores, and a simple count of AIS-coded injuries were added in turn to a base model using age and gender.

Results

A total of 20,813 patients survived to 12 months and had at least one functional outcome recorded, representing 85% follow-up. Predictions using the base model varied substantially across outcome measures. Irrespective of the method used to classify the severity of injury, adding injury severity to the model significantly, but only slightly improved model fit. Across the outcomes evaluated, no method of injury severity assessment consistently outperformed any other.

Conclusions

Anatomical injury is a predictor of trauma outcome. However, injury severity as described by the FCI does not consistently improve discrimination, or even provide the best discrimination compared to AIS-based severity scores or a simple injury count.

INTRODUCTION

Estimating the disease burden arising from injury is vital for guiding prevention and management priorities. However, recovery trajectories following serious injury vary widely and may be influenced by many demographic, epidemiological and psychosocial factors.[1-9] The location, type and extent of anatomical injury have been identified as a predictor of outcomes,[1,2,4,6-8,10,11] but in a number of studies anatomical injury has explained only a small proportion of outcome variability.[10-12] The Abbreviated Injury Scale (AIS) [13] provides a widely-used codeset for classifying anatomical injury, although the severity assessments contained within each AIS code are known to be biased towards mortality risk.[13,14] The predictive Functional Capacity Index (FCI) [15,16] was developed to predict functional outcomes in trauma survivors 12 months after injury.[13,17] Revised and appended to the 2008 AIS (AIS08),[13] the FCI may provide for injury burden estimates using AIS data routinely collected in trauma registries.[12,18]

A recent study demonstrated that the severity levels assigned within the FCI agreed more closely with assessed 12-month outcomes than AIS severity levels.[12] However, agreement was only 'slight'[19], even after excluding a majority of patients on the basis of age, multi-trauma or the presence of comorbidity. Also, beyond considering the worst FCI severity assigned to a patient's injuries,[16,20] no methods exist for accounting for multiple injuries (whether predicted to be disabling or not) when using the FCI.[9]

This study aimed to determine the extent to which AIS-based and FCI-based scores are able to add to a simple predictive model of 12-month functional outcomes in a severely injured population. A secondary aim of the study was to explore and evaluate potential methods of using FCI scores in instances where patients have sustained multiple injuries.

METHODS

Patients and source

This study used data from severely injured adult trauma patients in the Australian state of Victoria. Patients were drawn from the Victorian State Trauma Registry (VSTR), a well-established population-based registry collecting data on hospitalised major trauma.[21] All Victorian hospitals receiving trauma submit data to the VSTR; complete inclusion and exclusion criteria are published elsewhere.[22] The dataset included patients sustaining blunt or penetrating trauma between January 2007 and June 2015. Patients aged less than 18 years, or sustaining burn or asphyxia injury were excluded as the FCI was not designed for these patients.[16]

The VSTR collects cross-sectional data at several points following injury via standardised telephone interview of survivors to discharge (or their carers).[6] Two subgroups of patients receive this follow-up. The first of these are patients meeting Victorian major trauma criteria.[22] The other subgroup are co-included in the Victorian Orthopaedic Trauma Outcomes Registry (VOTOR),[23] which collects data on orthopaedic trauma admitted for more than 24 hours to one of four large sentinel hospitals. For this study, 12 month follow-up data were used.

Outcome measures used

The outcomes of interest were:

1) Glasgow Outcome Scale - Extended (GOS-E).[22,24] This eight-point hierarchical scale has been validated for use in general trauma populations.[24] A score of 5 or higher is representative of 'independent living', [25] and this dichotomisation was used.

2) Return to work status. Patients who had been working prior to the injury event were dichotomised depending on whether or not they resumed working.

3) The EQ-5D-3L (EQ-5D).[26] This generic measure of health status, including five items (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) measured on a 3-level scale (no, some or severe problems) has been recommended for evaluating trauma patients.[27] Responses to each item were dichotomised into ‘‘no problems’’ and ‘‘some/severe problems’’.[28]

Injury severity scores used

Nearly 10% of the AIS08's 1,999 codes either do not have FCI severities assigned (88 codes - 52 relating to blunt or penetrating injury), or represent minor superficial injuries with both AIS level 1 and FCI level 5 (90 codes). These injuries were excluded from analysis. They are listed in Supplement 1, along with their incidence in the study dataset. In order to evaluate the FCI as a single tool, a pragmatic approach was used to compare overall discrimination using FCI-based and AIS-based summary scores. This necessitated the development of two scores utilising FCI-based severities; their rationale and structure are described in Supplement 2. The following scores were employed:

1) Three well-established AIS-based summary scores:

i) Maximum AIS severity (MAIS);[29]

ii) Injury Severity Score (ISS);[30] and

iii) New Injury Severity Score (NISS).[31]

2) One established, and two novel, FCI-based summary scores:

i) Worst FCI;[9,16,20]

ii) Functional Capacity Additive Score (FCAS), a novel score which adjusts and adds the FCI severity levels of up to three worst injuries; and

iii) Functional Capacity Quadratic Score (FCQS), a novel score which adjusts and adds the squared FCI severity levels of up to three worst injuries in a similar manner to the NISS.

3) The total number of injuries (AIS codes) sustained. This functioned as an additional summary score independent of AIS or FCI severities.

Data analysis

Logistic regression was employed to test the predictive capacity of injury summary scores for each outcome. A split dataset approach using a 2:1 ratio was used, randomising cases to the 'training' dataset used to develop the model or the 'testing' dataset used to validate it. Predictors were not categorised, to avoid statistical inefficiency and loss of predictive power.[32]

The base model used only age [1-3,6,7,9] and gender,[2,7,9,25] which are well-recognised and universally comparable predictors of trauma outcome. Patient age was not restricted, although sensitivity analyses with restricted age groups [15,16] were performed and reported as supplementary data (Supplement 3).

Injury severity measures were added to the base model in turn. Ungrouped standardised Pearson chi-square tests were used to assess calibration in preference to the Hosmer-Lemeshow test, as many of the quantiles had substantial numbers of ties.[33,34] Discrimination was assessed using the area under the receiver operating characteristic (ROC) curve (AUC); Gönen's method was used to compare AUCs including injury severity measures.[35] Proportions were assessed with chi square testing, including evaluation of standardised residuals.

All analyses were performed using Stata Version 14.0 (StataCorp, College Station, Texas). A P-value less than 0.005 was considered significant;[36] confidence intervals and standardised residuals were reported at

the 99% level. Ethical approval for the study was granted by the Monash University Human Research Ethics Committee.

RESULTS

Derivation and description of the dataset

A total of 28,793 adult patients with blunt or penetrating trauma were retrieved from the VSTR (Fig 1). Loss to follow-up was low; of the 26,077 patients who survived to hospital discharge, 85.3% had a known 12-month outcome (Fig 1).

Surviving patients with valid AIS and FCI data available are summarised in Table 1. Patients lost to follow-up were more likely to be aged less than 45 years, and to reside in a socioeconomically disadvantaged area, based on the 2011 Australian census.[37] They sustained fewer falls injuries, and a higher proportion of penetrating (piercing, cutting or gunshot) and intentional injury or injury of unknown intent. These patients were also more likely to sustain injuries to 'other' regions such as the chest or abdomen, or to sustain only injuries of FCI level 5.

Splitting of patients into training and testing datasets returned comparable datasets (Table 1); the training dataset comprised 13,885 patients, and the testing dataset 6,928 patients out of a total of 20,813 patients with at least one outcome recorded. Of these, (n=20,777; 99.8%) had GOS-E scores recorded, and most (18,238; 87.6%) had one or more EQ-5D items recorded. In total, 12,283 patients (59.0%) had been working prior to injury; of these, almost all (12,151; 98.9%) had return to work status recorded, with the most common pre-injury occupation groups being tradespersons (28%), professional workers (14%) and clerical or service staff (12%). Most patients sustained multiple injuries (across one or more body regions); only 3,468 of 20,813 patients with 12-month outcomes (16.7%) sustained a single, non-superficial injury.

Table 1. Demographics and injury severity of surviving Victorian major and severe orthopaedic trauma patients with valid AIS and FCI data available. Percentages for each breakdown are shown in brackets, and do not always add to 100% due to rounding.

	Training dataset	Testing dataset	Lost to follow-up	Total
Total patients	**13,885**	**6,928**	**3,834**	**24,647**
Gender				
Male	9,661 (70)	4,740 (68)	2,734 (71)	**17,135 (70)**
Female	4,224 (30)	2,188 (32)	1,100 (29)	**7,512 (30)**
Age group				
18-24 years	1,930 (14)	1,002 (14)	712 (19)	**3,644 (15)**
25-34 years	2,003 (14)	1,000 (14)	840 (22)	**3,843 (16)**
35-44 years	1,972 (14)	951 (14)	669 (17)	**3,592 (15)**
45-54 years	2,053 (15)	961 (14)	493 (13)	**3,507 (14)**
55-64 years	1,774 (13)	906 (13)	364 (9)	**3,044 (12)**
65-74 years	1,594 (11)	784 (11)	265 (7)	**2,643 (11)**
75-84 years	1,606 (12)	849 (12)	299 (8)	**2,754 (11)**
85+ years	953 (7)	475 (7)	192 (5)	**1,620 (7)**
Comorbidity present				
Healthy (CCI [a] = 0)	9,385 (68)	4,698 (68)	2,657 (69)	**16,740 (68)**
Comorbidity (CCI [a] >0)	4,500 (32)	2,230 (32)	1,177 (31)	**7,907 (32)**
IRSAD decile [b]				
1st quintile (most disadvantaged)	1,860 (13)	900 (13)	653 (17)	**3,413 (14)**
2nd quintile	1,921 (14)	999 (14)	536 (14)	**3,456 (14)**
3rd quintile	2,752 (20)	1,421 (21)	742 (19)	**4,915 (20)**
4th quintile	3,093 (22)	1,521 (22)	697 (18)	**5,311 (22)**
5th quintile (most advantaged)	3,940 (28)	1,905 (28)	966 (25)	**6,811 (28)**
Unknown	319 (2)	182 (3)	240 (6)	**741 (3)**
Intent of injury				
Unintentional	12,668 (91)	6,335 (91)	3,096 (81)	**22,099 (90)**
Intentional (assault or self-harm)	1,030 (7)	507 (7)	632 (16)	**2,169 (9)**
Other or unspecified intent	187 (1)	86 (1)	106 (3)	**379 (2)**
Compensability of injuries				
Compensable	6,075 (44)	3,031 (44)	1,589 (41)	**10,695 (43)**
Non-compensable	7,810 (56)	3,897 (56)	2,245 (59)	**13,952 (57)**
Mechanism of Injury				
Occupant in motor vehicle	3,159 (23)	1,612 (23)	887 (23)	**5,658 (23)**
Other transport-related	3,490 (25)	1,712 (25)	828 (22)	**6,030 (24)**
Fall ≤ 1m	3,344 (24)	1,746 (25)	803 (21)	**5,893 (24)**
Fall > 1m	1,788 (13)	841 (12)	406 (11)	**3,035 (12)**
Piercing, cutting or gunshot	414 (3)	193 (3)	294 (8)	**901 (4)**
Other mechanism	1,690 (12)	824 (12)	616 (16)	**3,130 (13)**
Reason for VSTR [c] inclusion				
Major trauma	4,503 (32)	2,232 (32)	1,718 (45)	**8,453 (34)**

	Training dataset	Testing dataset	Lost to follow-up	Total
Orthopaedic trauma	4,300 (31)	2,147 (31)	1,124 (29)	**7,571 (31)**
Major and orthopaedic trauma	5,082 (37)	2,549 (37)	992 (26)	**8,623 (35)**
Maximum AIS score				
1	48 (0)	21 (0)	31 (1)	**100 (0)**
2	3,572 (26)	1,788 (26)	1,061 (28)	**6,421 (26)**
3	6,206 (45)	3,063 (44)	1,716 (45)	**10,985 (45)**
4	2,713 (20)	1,350 (19)	687 (18)	**4,750 (19)**
5	1,340 (10)	705 (10)	337 (9)	**2,382 (10)**
6	6 (0)	1 (0)	2 (0)	**9 (0)**
Worst FCI score				
5 (best outcome)	7,372 (53)	3,667 (53)	2,300 (60)	**13,339 (54)**
4	2,308 (17)	1,191 (17)	555 (14)	**4,054 (16)**
3	1,079 (8)	494 (7)	265 (7)	**1,838 (7)**
2	1,421 (10)	736 (11)	298 (8)	**2,455 (10)**
1 (worst outcome)	1,705 (12)	840 (12)	416 (11)	**2,961 (12)**
ISS grouping [d]				
1 - 8	2,710 (20)	1,363 (20)	812 (21)	**4,885 (20)**
9 - 12	2,741 (20)	1,342 (19)	762 (20)	**4,845 (20)**
13 - 14	2,283 (16)	1,109 (16)	831 (22)	**4,223 (17)**
16 - 19	2,706 (19)	1,402 (20)	657 (17)	**4,765 (19)**
20 - 24	1,283 (9)	620 (9)	265 (7)	**2,168 (9)**
25 - 38	1,934 (14)	971 (14)	448 (12)	**3,353 (14)**
41 - 48	143 (1)	76 (1)	34 (1)	**253 (1)**
50 - 75	85 (1)	45 (1)	25 (1)	**155 (1)**
Body regions injured				
Head only	1,446 (10)	729 (11)	371 (10)	**2,546 (10)**
Head + spinal cord	269 (2)	125 (2)	42 (1)	**442 (2)**
Head + other	4,098 (30)	2,065 (30)	1,067 (28)	**7,230 (29)**
Other spinal cord	603 (4)	290 (4)	150 (4)	**1,043 (4)**
Orthopaedic only	3,365 (24)	1,688 (24)	865 (23)	**5,918 (24)**
Orthopaedic + other	3,174 (23)	1,569 (23)	878 (23)	**5,621 (23)**
Other	930 (7)	462 (7)	455 (12)	**1,847 (7)**

[a] CCI - Charlson Comorbidity Index.

[b] IRSAD - Index of Relative Socio-economic Advantage and Disadvantage.

[c] VSTR - Victorian State Trauma Registry.

[d] ISS - Injury Severity Score. Not all ISS values (such as 15, 39, 40 and 49) are obtainable, due to the construction of the ISS.

Training dataset

The base model of age and gender alone varied substantially in predicting functional outcomes; predictions for return to work and the EQ-5D items of pain/discomfort and anxiety/depression were little better than chance, while predictions of GOS-E had an AUC of 0.762 (Table 2). Irrespective of the summary score used, adding injury to the model significantly improved model fit; the sole exception was the addition of NISS in the prediction of EQ-5D pain/discomfort (Table 2). Models using the FCAS produced the highest AUC for return to work and the EQ-5D mobility and usual activities items, and models using the FCQS the highest AUC for the GOS-E and EQ-5D personal care item. However, models using the simple count of the number of injuries produced the highest AUC for the EQ-5D pain/discomfort and anxiety/depression items (Table 2). Models only exceeded an AUC of 0.70 for three outcomes - GOS-E and the EQ-5D mobility and personal care items - and were never higher than 0.60 for the EQ-5D pain/discomfort and anxiety/depression items. Models predicting GOS-E were not well-calibrated, but models predicting other outcomes were generally well-calibrated (Table 2).

No method of injury severity assessment consistently outperformed any other (Table 2). The variability in discrimination across each outcome was often small - for example, all of the injury-adjusted models predicting the EQ-5D personal care item varied between 0.717 and 0.727 (Table 2; Fig 2). There was no significant difference between any of the injury-adjusted models for two of the physical outcomes (GOS-E and the personal care item of the EQ-5D) and one psychosocial outcome (the pain/discomfort item of the EQ-5D). On two of the outcome measures (the mobility and usual activities items of the EQ-5D), all AIS-based models performed significantly worse than the highest (FCI-based) model. However amongst models predicting the EQ-5D anxiety/depression item, the second-highest discrimination (after the number of injuries) was the model containing the ISS (Table 2).

Testing dataset

When the same models were fitted to the testing dataset, results were similar (Table 3). All injury scores improved model fit for all outcomes, with the exception of the MAIS for predicting the EQ-5D usual activities item, or any AIS-based score for the EQ-5D pain/discomfort item. Again, no single method of injury adjustment consistently produced higher discrimination. The number of injuries sustained produced the highest AUC when predicting return to work or the EQ-5D pain/discomfort and anxiety/depression items; the FCAS for the EQ-5D mobility and usual activities items; the FCQS for the EQ-5D personal care item; and the MAIS the highest AUC when predicting GOS-E.

Table 2. Discrimination and calibration of models in the training dataset (total n = 13,885 patients). For each outcome, likelihood ratios are compared to the base model of age and gender. The highest area under the ROC curve (AUC) for each outcome is indicated with an asterisk; models including a measure of anatomical injury giving an AUC significantly lower than this (with a chi-square statistic of 7.879 equivalent to a P-value of 0.005) are indicated by an obelisk.

Model outcome	Area under ROC curve (99% CI)	Ungrouped Pearson X^2 statistic (P-value)	LR test (P-value) [a]	X^2 difference to highest AUC [b] (P-value)
GOS-E [b] outcome	**(n=13,866)**			
Age and gender	0.762 (0.748, 0.777)	15,492.2 (<0.0001)	-	-
Age, gender and no. of injuries	0.769 (0.755, 0.783)	15,265.0 (<0.0001)	64.70 (<0.0001)	4.64 (0.031)
Age, gender and MAIS [d]	0.779 (0.766, 0.793)	14,865.7 (0.0005)	270.73 (<0.0001)	0.53 (0.467)
Age, gender and ISS [e]	0.781 (0.768, 0.794)	14,934.4 (0.0007)	257.42 (<0.0001)	0.26 (0.611)
Age, gender and NISS [f]	0.779 (0.766, 0.793)	14,922.7 (0.0007)	225.86 (<0.0001)	0.55 (0.458)
Age, gender and worst FCI [g]	0.779 (0.766, 0.792)	14,647.4 (0.001)	232.38 (<0.0001)	0.58 (0.446)
Age, gender and FCAS [h]	0.778 (0.765, 0.791)	14,831.9 (<0.0001)	201.51 (<0.0001)	0.85 (0.358)
Age, gender and FCQS [i]	0.785 (0.772, 0.798) *	14,809.0 (0.005)	308.62 (<0.0001)	-
Return to work outcome	**(n=8,132)**			
Age and gender	0.527 (0.509, 0.545)	8,132.7 (0.987)	-	-
Age, gender and no. of injuries	0.632 (0.614, 0.650)	8,126.0 (0.922)	382.79 (<0.0001)	3.30 (0.069)
Age, gender and MAIS [d]	0.601 (0.583, 0.618)	8,125.6 (0.912)	222.37 (<0.0001)	25.27 (<0.0001) †
Age, gender and ISS [e]	0.627 (0.610, 0.645)	8,149.1 (0.789)	373.85 (<0.0001)	5.25 (0.021)
Age, gender and NISS [f]	0.622 (0.604, 0.639)	8,118.0 (0.817)	321.79 (<0.0001)	8.24 (0.004) †
Age, gender and worst FCI [g]	0.637 (0.619, 0.655)	8,118.7 (0.039)	373.10 (<0.0001)	1.72 (0.190)
Age, gender and FCAS [h]	0.650 (0.632, 0.667) *	8,118.8 (0.542)	491.36 (<0.0001)	-
Age, gender and FCQS [i]	0.643 (0.625, 0.661)	8,106.5 (0.661)	437.77 (<0.0001)	0.22 (0.640)
EQ-5D mobility outcome	**(n=12,200)**			
Age and gender	0.683 (0.670, 0.696)	12,264.5 (0.485)	-	-
Age, gender and no. of injuries	0.702 (0.689, 0.714)	12,271.2 (0.471)	202.02 (<0.0001)	12.39 (0.004) †
Age, gender and MAIS [d]	0.688 (0.675, 0.700)	12,258.9 (0.419)	54.67 (<0.0001)	31.30 (<0.0001) †
Age, gender and ISS [e]	0.694 (0.681, 0.706)	12,267.1 (0.466)	116.37 (<0.0001)	21.97 (<0.0001) †
Age, gender and NISS [f]	0.690 (0.678, 0.703)	12,255.7 (0.528)	73.55 (<0.0001)	26.97 (<0.0001) †
Age, gender and worst FCI [g]	0.720 (0.708, 0.732)	12,214.2 (0.737)	462.71 (<0.0001)	0.75 (0.386)
Age, gender and FCAS [h]	0.725 (0.713, 0.737) *	12,253.6 (0.018)	544.76 (<0.0001)	-
Age, gender and FCQS [i]	0.720 (0.708, 0.732)	12,275.4 (0.496)	454.96 (<0.0001)	0.34 (0.561)
EQ-5D personal care outcome	**(n=12,196)**			
Age and gender	0.710 (0.696, 0.725)	12,511.0 (0.068)	-	-
Age, gender and no. of injuries	0.722 (0.709, 0.736)	12,462.5 (0.129)	121.95 (<0.0001)	0.27 (0.600)
Age, gender and MAIS [d]	0.717 (0.703, 0.731)	12,410.1 (0.132)	74.32 (<0.0001)	1.58 (0.208)
Age, gender and ISS [e]	0.721 (0.707, 0.735)	12,408.8 (0.210)	108.61 (<0.0001)	0.56 (0.455)
Age, gender and NISS [f]	0.718 (0.704, 0.732)	12,415.6 (0.187)	83.34 (<0.0001)	1.10 (0.293)
Age, gender and worst FCI [g]	0.723 (0.709, 0.737)	12,348.9 (0.209)	140.17 (<0.0001)	0.19 (0.666)
Age, gender and FCAS [h]	0.725 (0.711, 0.739)	12,372.6 (0.049)	161.66 (<0.0001)	0.02 (0.875)
Age, gender and FCQS [i]	0.727 (0.713, 0.740) *	12,381.5 (0.313)	180.47 (<0.0001)	-
EQ-5D usual activities outcome	**(n=12,186)**			

Model outcome	Area under ROC curve (99% CI)	Ungrouped Pearson X^2 statistic (P-value)	LR test (P-value) [a]	X^2 difference to highest AUC [b] (P-value)
Age and gender	0.598 (0.585, 0.611)	12,173.4 (0.663)	-	-
Age, gender and no. of injuries	0.632 (0.619, 0.645)	12,190.1 (0.923)	269.93 (<0.0001)	0.82 (0.365)
Age, gender and MAIS [d]	0.601 (0.588, 0.614)	12,173.9 (0.642)	24.19 (<0.0001)	27.27 (<0.0001) †
Age, gender and ISS [e]	0.613 (0.600, 0.626)	12,176.9 (0.804)	105.00 (<0.0001)	12.95 (0.0003) †
Age, gender and NISS [f]	0.609 (0.596, 0.622)	12,177.5 (0.806)	76.38 (<0.0001)	16.83 (0.0004) †
Age, gender and worst FCI [g]	0.630 (0.617, 0.643)	12,201.1 (0.005)	223.25 (<0.0001)	1.41 (0.235)
Age, gender and FCAS [h]	0.639 (0.626, 0.651) *	12,213.2 (0.171)	315.56 (<0.0001)	-
Age, gender and FCQS [i]	0.632 (0.619, 0.645)	12,214.5 (0.495)	246.46 (<0.0001)	0.38 (0.535)
EQ-5D pain/discomfort outcome	**(n=12,109)**			
Age and gender	0.530 (0.517, 0.544)	12,108.9 (0.959)	-	-
Age, gender and no. of injuries	0.584 (0.571, 0.597) *	12,121.6 (0.505)	234.90 (<0.0001)	-
Age, gender and MAIS [d]	0.540 (0.527, 0.554)	12,109.0 (0.998)	13.13 (0.0003)	35.47 (<0.0001) †
Age, gender and ISS [e]	0.538 (0.525, 0.551)	12,108.8 (0.949)	15.27 (0.0001)	39.30 (<0.0001) †
Age, gender and NISS [f]	0.535 (0.521, 0.548)	12,108.9 (0.926)	2.72 (0.099)	45.31 (<0.0001) †
Age, gender and worst FCI [g]	0.549 (0.535, 0.562)	12,109.6 (0.953)	36.58 (<0.0001)	23.04 (<0.0001) †
Age, gender and FCAS [h]	0.570 (0.557, 0.584)	12,114.6 (0.765)	105.11 (<0.0001)	3.62 (0.057)
Age, gender and FCQS [i]	0.554 (0.541, 0.568)	12,111.1 (0.552)	39.84 (<0.0001)	8.32 (0.004) †
EQ-5D anxiety/depression outcome	**(n=12,082)**			
Age and gender	0.544 (0.530, 0.558)	12,080.9 (0.944)	-	-
Age, gender and no. of injuries	0.574 (0.560, 0.587) *	12,080.9 (0.967)	120.04 (<0.0001)	-
Age, gender and MAIS [d]	0.555 (0.541, 0.569)	12,081.9 (0.996)	37.56 (<0.0001)	5.81 (0.016)
Age, gender and ISS [e]	0.568 (0.554, 0.582)	12,082.5 (0.985)	99.22 (<0.0001)	0.50 (0.480)
Age, gender and NISS [f]	0.566 (0.552, 0.579)	12,082.6 (0.982)	83.27 (<0.0001)	1.07 (0.301)
Age, gender and worst FCI [g]	0.557 (0.543, 0.570)	12,080.7 (0.458)	40.02 (<0.0001)	5.01 (0.025)
Age, gender and FCAS [h]	0.567 (0.553, 0.581)	12,080.9 (0.893)	87.11 (<0.0001)	0.78 (0.376)
Age, gender and FCQS [i]	0.561 (0.547, 0.574)	12,081.1 (0.965)	64.18 (<0.0001)	1.45 (0.229)

[a] LR - Likelihood ratio test; compared to model with age and gender only.

[b] AUC - Area under the receiver operating characteristic (ROC) curve.

[c] GOS-E - Extended Glasgow Outcome Scale.

[d] MAIS - Maximum 2008 Abbreviated Injury Scale severity.

[e] ISS - Injury Severity Score.

[f] NISS - New Injury Severity Score.

[g] Worst FCI - Worst 2008 predictive Functional Capacity Index score.

[h] FCAS - Functional Capacity Additive Score (based on three worst FCI scores).

[i] FCQS - Functional Capacity Quadratic Score (based on up to three worst FCI scores).

Table 3. Discrimination and calibration of models in the testing dataset (total n = 6,928 patients). For each outcome, likelihood ratios are compared to the base model of age and gender. The highest area under the ROC curve (AUC) for each outcome is indicated with an asterisk.

Model outcome	Area under ROC curve (99% CI)	Ungrouped Pearson X^2 statistic (P-value)	LR test (P-value) [a]
GOS-E [b] outcome	**(n=6,911)**		
Age and gender	0.762 (0.742, 0.782)	7,565.0 (0.002)	-
Age, gender and no. of injuries	0.765 (0.746, 0.785)	7,504.5 (0.003)	12.78 (0.0004)
Age, gender and MAIS [d]	0.780 (0.761, 0.799) *	7,402.3 (0.017)	143.14 (<0.0001)
Age, gender and ISS [e]	0.778 (0.759, 0.796)	7,462.2 (0.013)	114.84 (<0.0001)
Age, gender and NISS [f]	0.778 (0.760, 0.797)	7,445.6 (0.016)	112.29 (<0.0001)
Age, gender and worst FCI [g]	0.776 (0.757, 0.795)	7,343.6 (0.012)	95.08 (<0.0001)
Age, gender and FCAS [h]	0.774 (0.755, 0.793)	7,461.4 (0.0001)	70.46 (<0.0001)
Age, gender and FCQS [i]	0.779 (0.761, 0.798)	7,449.1 (0.019)	116.07 (<0.0001)
Return to work outcome	**(n=4,019)**		
Age and gender	0.538 (0.512, 0.564)	4,018.8 (0.995)	-
Age, gender and no. of injuries	0.638 (0.612, 0.663)	4,012.8 (0.894)	213.62 (<0.0001)
Age, gender and MAIS [d]	0.605 (0.580, 0.630)	4,019.6 (0.990)	113.25 (<0.0001)
Age, gender and ISS [e]	0.630 (0.605, 0.655)	4,022.8 (0.936)	184.28 (<0.0001)
Age, gender and NISS [f]	0.632 (0.607, 0.657)	4,007.2 (0.800)	176.64 (<0.0001)
Age, gender and worst FCI [g]	0.629 (0.604, 0.654)	4,017.8 (0.852)	164.88 (<0.0001)
Age, gender and FCAS [h]	0.638 (0.613, 0.662) *	4,000.2 (0.045)	185.62 (<0.0001)
Age, gender and FCQS [i]	0.634 (0.609, 0.659)	4,006.6 (0.767)	168.94 (<0.0001)
EQ-5D mobility outcome	**(n=6,026)**		
Age and gender	0.702 (0.684, 0.720)	6,053.4 (0.703)	-
Age, gender and no. of injuries	0.711 (0.694, 0.729)	6,045.8 (0.779)	44.69 (<0.0001)
Age, gender and MAIS [d]	0.705 (0.688, 0.723)	6,043.8 (0.751)	26.08 (<0.0001)
Age, gender and ISS [e]	0.710 (0.691, 0.727)	6,048.1 (0.750)	47.25 (<0.0001)
Age, gender and NISS [f]	0.708 (0.690, 0.725)	6,046.1 (0.764)	32.07 (<0.0001)
Age, gender and worst FCI [g]	0.728 (0.711, 0.745)	6,031.2 (0.890)	171.34 (<0.0001)
Age, gender and FCAS [h]	0.731 (0.714, 0.748) *	6,041.1 (0.474)	202.29 (<0.0001)
Age, gender and FCQS [i]	0.729 (0.712, 0.746)	6,040.7 (0.859)	185.17 (<0.0001)
EQ-5D personal care outcome	**(n=6,021)**		
Age and gender	0.727 (0.707, 0.747)	6,171.6 (0.267)	-
Age, gender and no. of injuries	0.733 (0.713, 0.752)	6,124.5 (0.426)	22.57 (<0.0001)
Age, gender and MAIS [d]	0.733 (0.713, 0.753)	6,166.4 (0.199)	39.69 (<0.0001)
Age, gender and ISS [e]	0.737 (0.717, 0.756)	6,158.9 (0.303)	56.30 (<0.0001)
Age, gender and NISS [f]	0.734 (0.715, 0.754)	6,151.5 (0.316)	38.17 (<0.0001)
Age, gender and worst FCI [g]	0.739 (0.720, 0.758)	6,095.2 (0.457)	65.67 (<0.0001)
Age, gender and FCAS [h]	0.738 (0.719, 0.758)	6,117.0 (0.212)	59.05 (<0.0001)
Age, gender and FCQS [i]	0.740 (0.721, 0.760) *	6,123.3 (0.473)	72.56 (<0.0001)
EQ-5D usual activities outcome	**(n=6,020)**		
Age and gender	0.619 (0.601, 0.638)	6,014.7 (0.842)	-
Age, gender and no. of injuries	0.632 (0.614, 0.651)	6,014.0 (0.859)	78.73 (<0.0001)

Model outcome	Area under ROC curve (99% CI)	Ungrouped Pearson X^2 statistic (P-value)	LR test (P-value) [a]
Age, gender and MAIS [d]	0.621 (0.602, 0.639)	6,014.8 (0.811)	9.63 (0.0019)
Age, gender and ISS [e]	0.625 (0.607, 0.644)	6,013.7 (0.830)	37.42 (<0.0001)
Age, gender and NISS [f]	0.623 (0.604, 0.641)	6,014.8 (0.852)	24.14 (<0.0001)
Age, gender and worst FCI [g]	0.639 (0.620, 0.657)	6,028.2 (0.204)	81.75 (<0.0001)
Age, gender and FCAS [h]	0.642 (0.624, 0.660) *	6,024.3 (0.522)	110.65 (<0.0001)
Age, gender and FCQS [i]	0.640 (0.622, 0.659)	6,025.2 (0.876)	89.40 (<0.0001)
EQ-5D pain/discomfort outcome	**(n=5,978)**		
Age and gender	0.534 (0.515, 0.553)	5,978.0 (0.904)	-
Age, gender and no. of injuries	0.568 (0.549, 0.587) *	5,980.0 (0.827)	64.17 (<0.0001)
Age, gender and MAIS [d]	0.539 (0.519, 0.558)	5,977.9 (0.979)	4.08 (0.0433)
Age, gender and ISS [e]	0.537 (0.518, 0.556)	5,977.8 (0.938)	6.53 (0.0106)
Age, gender and NISS [f]	0.535 (0.516, 0.554)	5,978.0 (0.834)	0.22 (0.6400)
Age, gender and worst FCI [g]	0.548 (0.528, 0.567)	5,978.3 (0.967)	15.43 (0.0001)
Age, gender and FCAS [h]	0.564 (0.549, 0.583)	5,980.0 (0.865)	43.45 (<0.0001)
Age, gender and FCQS [i]	0.554 (0.535, 0.573)	5,979.2 (0.712)	20.89 (<0.0001)
EQ-5D anxiety/depression outcome	**(n=5,980)**		
Age and gender	0.543 (0.523, 0.563)	5,980.7 (0.957)	-
Age, gender and no. of injuries	0.562 (0.542, 0.582) *	5,981.3 (0.939)	27.87 (<0.0001)
Age, gender and MAIS [d]	0.551 (0.531, 0.570)	5,980.7 (0.960)	11.18 (0.0008)
Age, gender and ISS [e]	0.556 (0.536, 0.575)	5,981.5 (0.929)	25.10 (<0.0001)
Age, gender and NISS [f]	0.558 (0.539, 0.578)	5,980.8 (0.962)	23.07 (<0.0001)
Age, gender and worst FCI [g]	0.553 (0.533, 0.573)	5,981.3 (0.727)	13.02 (0.0003)
Age, gender and FCAS [h]	0.560 (0.540, 0.579)	5,981.2 (0.493)	23.14 (<0.0001)
Age, gender and FCQS [i]	0.557 (0.537, 0.576)	5,980.9 (0.956)	18.72 (<0.0001)

[a] LR - Likelihood ratio test; compared to model with age and gender only.

[b] AUC - Area under the receiver operating characteristic (ROC) curve.

[c] GOS-E - Extended Glasgow Outcome Scale.

[d] MAIS - Maximum 2008 Abbreviated Injury Scale severity.

[e] ISS - Injury Severity Score.

[f] NISS - New Injury Severity Score.

[g] Worst FCI - Worst 2008 predictive Functional Capacity Index score.

[h] FCAS - Functional Capacity Additive Score (based on three worst FCI scores).

[i] FCQS - Functional Capacity Quadratic Score (based on up to three worst FCI scores).

Models predicting GOS-E were again poorly calibrated, but acceptable calibration was observed for other outcomes (Table 3; Fig 3). While middle-range prediction predictions tracked close to the line of best fit, calibration was poorer at scale extremes. Both the base model, and the models incorporating AIS-based injury adjustment, tended to over-predict outcomes at lower prediction quantiles, and under-predicted all but GOS-E and the EQ-5D personal care item at higher quantiles. In contrast, models incorporating FCI-based injury adjustment tended to under-predict outcomes at lower quantiles, and over-predicted outcomes at higher prediction quantiles.

DISCUSSION

The authors who first presented the FCI stated 20 years ago that, "the FCI must be empirically validated across the full spectrum of injury type and severity... An important aspect of the validation will be the comparison of the FCI with widely accepted performance based and self-reported measures of function".[16] In the present study, the AIS and FCI often performed similarly in improving the prediction of outcomes over models using age and gender alone. The FCI was developed to provide outcomes-weighted severities as an alternative to the mortality-weighted severities in the AIS codeset. In this respect, the FCI is unlikely to be fit for purpose as a global outcome prediction tool.

Previous studies have found associations between anatomical injury and a range of physical outcomes.[1-7,25,28,38] While some studies have found that the ISS is independently associated with functional outcomes,[1,5] others have found that injuries to particular body regions or the presence of multi-trauma contribute variously to different outcome measures.[1-7,38] In the present study, models containing the AIS-based, mortality-weighted ISS significantly improved model performance for all but one outcome, and returned a higher AUC than all FCI-based models in predicting the anxiety/depression outcome of the EQ-5D (Table 2).

Although models incorporating injury severity performed relatively well on physical measures, for psychosocial measures (the pain/discomfort and anxiety/depression components of the EQ-5D) they performed little better than chance; this is in keeping with previous findings.[25] Pain and psychosocial outcomes were specifically excluded from the dimensions of function covered by the FCI;[15] this has previously been criticised.[27,39] However, the types of injury - and the non-injury predictors - which contribute to physical outcomes are known to differ from those which contribute to psychosocial outcomes.[4,5,38] As

such, it is unlikely that scores using a single severity level for 'outcome' (such as that offered by the AIS or FCI)[13] will satisfactorily add to models predicting both functional (physical) outcomes and quality of life.

In this context, the performance of the simple number of coded injuries in predicting the outcome measures assessed - particularly the EQ-5D psychosocial dimensions - is unsurprising. The number of injuries has previously been associated with outcome across all dimensions of the EQ-5D.[28] In the present study, the model containing the number of injuries outperformed AIS-based models for all outcomes except GOS-E, although FCI-based models were consistently higher for physical outcome measures including return to work. The present study created two new FCI-based summary scores in order to validate the FCI for these patients. However, none of the AIS- or FCI-based summary scores utilised severity data for more than three injuries. Consequently, there may be better ways to utilise FCI severities in the presence of multiple injuries.

The FCI was designed specifically to provide a function-weighted alternative to the mortality-weighted severities comprising the AIS,[16] and outperformed the AIS in an earlier study assessing agreement between anatomical injury severity and GOS-E.[12] As such, it is unclear why FCI-based scores often provided only marginal gains over AIS-based scores in the present study even when their severities were used in similar ways (as with the NISS and FCQS). The pFCI08's developers used a standard gamble methodology to derive FCI severities.[20] However, these rely on accurate clinician descriptions of the expected functional outcome of each injury, and as such may be unsuitable for a highly specific classification system such as the AIS. Previous studies have found greater variability in FCI predictions for injuries to the head, lower limb and spine.[8,17] As a result, it is unsurprising that outcome predictions in a population with mixed major and orthopaedic trauma are less accurate than might be anticipated given the aims of the FCI. However, the exact extent to which anatomical injury predicts different functional outcomes - as estimated using several methods in the present study - remains unclear.

The present study used two novel methods for combining FCI scores in the presence of multiple injuries. This is essential to routine outcome prediction; in the present study the majority of patients sustained injuries to multiple body regions (Table 1). For all but one of the outcomes evaluated, models containing either the FCAS or FCQS generally recorded slightly (although not significantly) higher AUC than the single worst FCI, which was previously the recommended method.[9,16,20] In addition, the FCAS was the only score not to differ significantly from the best performing model (including the number of injuries) in predicting the

pain/discomfort outcome of the EQ-5D. As such, this study serves as a de facto validation of these summary scores.

Particular strengths of this study included the opt-off consent process and high follow-up rates recorded on the VSTR, and the inclusive trauma system which formed the setting for the study. A further strength is the inclusion of less severely injured orthopaedic trauma patients in addition to major trauma. Orthopaedic injuries have been found to account for the majority of years lived with disability amongst trauma patients admitted to hospital,[40] although many studies assessing trauma outcomes have focused on major trauma.[3-7,38] However, there were some limitations with the present study. Patients lost to follow-up differed from included cases in terms of gender, socio-economic status, mechanism and intent of injury. As such, there may be biases which affect the interpretation of the study's findings. However, these are likely to be minor given the comparatively small associations between assessed functional outcomes and both the FCI and AIS.[12]

Dichotomisation of assessed outcomes is appropriate,[32] and has been used for these outcomes.[25,27,28] However, it is possible that individual predictors may have greater or lesser effects at different levels of function. For example, gender may be poor at discriminating between GOS-E of 2 and 3, but effective at discriminating between GOS-E of 7 and 8. Similar effects may also be present across injuries of different types, or to different body regions. However, the evaluation of sub-groups of patients was outside the scope of the present study which sought to evaluate the overall performance of the FCI.

Similarly, AIS- and FCI-based scores are known to be ordinal, rather than continuous measures. However, ordinal logistic regression methods still assume a proportional variation between values. Given that this may not be the case, the techniques used were believed to be reasonable.

Other predictive factors such as education level, the presence of comorbidities, and the compensability of a patient's injuries have been shown to predict both physical and psychosocial outcomes.[38] Model performance may have improved with the addition of these variables. However, the intent of the study was not to develop optimal models for outcome prediction, but to assess the effects of different methods of categorising injury within such models.

CONCLUSION

Anatomical injury is a significant predictor of longer term functional, occupational and quality of life outcomes. Adding injury severity to a simple model improves the prediction of outcomes after serious injury. However, injury severity as described by the FCI does not consistently increase discrimination, or provide for the best discrimination, when compared to AIS-based severity scores or a simple count of the injuries sustained. In order to maximise their effectiveness, models predicting different aspects of physical or psychosocial recovery after severe trauma may require quite different representations of anatomical injury severity which may not be based on either AIS or FCI severities.

Acknowledgements

None.

Contributors

CSP, PAC and BJG are investigators of the project, contributed to the study design, reviewed the manuscript and approved the final version of the manuscript. CSP completed all analyses for the study and drafted the manuscript.

Funding

The VSTR is a Department of Health, State Government of Victoria and Transport Accident Commission (TAC) funded project. VOTOR is funded by the TAC via the Institute for Safety, Compensation and Recovery Research. BJG and PAC are supported by a Career Development Fellowship (GNT1048731), and a Practitioner Fellowship (ID 545926), from the NHMRC, respectively. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

None.

Ethics approval

Monash University HREC.

Data sharing statement

The data included in this project are not freely available. Requests for access to data from the participating datasets would need to be directed to the relevant data custodian, who can be contacted at susan.mclellan@monash.edu or at the following URLs: https://www.monash.edu/medicine/sphpm/vstorm/data-requests (VSTR), or https://www.monash.edu/medicine/sphpm/votor/data-requests (VOTOR).

REFERENCES

1. Schluter PJ, McClure RJ. Predicting functional capacity outcome 12 months after hospitalized injury. *ANZ J Surg* 2006;76(10):886-93.
2. Polinder S, Van Beeck EF, Essink-Bot ML, et al. Functional outcome at 2.5, 5, 9, and 24 months after injury in the Netherlands. *J Trauma* 2007;62(1):133-41.
3. Holtslag HR, Post MW, Lindeman E, et al. Long-term functional health status of severely injured patients. *Injury* 2007;38(3):280-89.
4. Holtslag HR, Van Beeck EF, Lindeman E, et al. Determinants of long-term functional consequences after major trauma. *J Trauma* 2007;62(4):919-27.
5. Ringburg AN, Polinder S, van Ierland MC, et al. Prevalence and prognostic factors of disability after major trauma. *J Trauma* 2011;70(4):916-22.
6. Gabbe BJ, Simpson PM, Sutherland AM, et al. Improved functional outcomes for major trauma patients in a regionalized, inclusive trauma system. *Ann Surg* 2012;255(6):1009-15.
7. Gabbe BJ, Simpson PM, Harrison JE, et al. Return to Work and Functional Outcomes After Major Trauma: Who Recovers, When, and How Well? *Ann Surg* 2016;263(4):623-32.
8. McMurry TL, Poplin GS, Crandall J. Functional recovery patterns in seriously injured automotive crash victims. *Traffic Inj Prev* 2016;17 Suppl 1:21-26.
9. Palmer CS, Cameron PA, Gabbe BJ. A review of the revised Functional Capacity Index as a predictor of 12 month outcomes following injury. *Injury* 2017;48(3):591-98.
10. Schluter PJ, Cameron CM, Purdie DM, et al. How well do anatomical-based injury severity scores predict health service use in the 12 months after injury? *Int J Inj Contr Saf Promot* 2005;12(4):241-46.
11. Schluter PJ, Neale R, Scott D, et al. Validating the Functional Capacity Index: A Comparison of Predicted versus Observed Total Body Scores. *J Trauma* 2005;58(2):259-63.
12. Palmer CS, Gabbe BJ, Cameron PA. Revised Functional Capacity Index as a predictor of outcome following injury. *Br J Surg* 2017;104(13):1874-83.
13. Gennarelli TA, Wodzin E, editors. *Abbreviated Injury Scale 2005 - Update 2008*. Barrington, IL: AAAM, 2008.
14. Huang LC, Marsh JC. AIS and threat to life. *22nd AAAM Ann Conf Proc* 1978;22:242-54.
15. MacKenzie EJ, Damiano AM, Ditunno JF, et al. Development of the Functional Capacity Index (FCI). Springfield, Virginia: National Highway Traffic Safety Administration, 1994.
16. MacKenzie EJ, Damiano A, Miller T, et al. The development of the Functional Capacity Index. *J Trauma* 1996;41(5):799-807.
17. Barnes J, Morris A. A study of impairing injuries in real world crashes using the Injury Impairment Scale (IIS) and the predicted Functional Capacity Index (PFCI-AIS). *Ann Adv Automot Med* 2009;53:195-205.
18. Lessons Learned Using the Functional Capacity Index (FCI). Measuring the Burden of Injury: 3rd Annual Conference Proceedings; 2000; Baltimore, Maryland.
19. Byrt T. How good is that agreement? *Epidemiology* 1996;7(5):561.
20. McMurry TL, Sherwood C, Poplin GS, et al. Implications of Functional Capacity Loss and Fatality for Vehicle Safety Prioritization. *Traffic Inj Prev* 2015;16(Sup2):S140-S45.
21. Gabbe BJ, Sutherland AM, Hart MJ, et al. Population-based capture of long-term functional and quality of life outcomes after major trauma: the experiences of the Victorian State Trauma Registry. *J Trauma* 2010;69(3):532-36.
22. Cameron PA, Gabbe BJ, Cooper DJ, et al. A statewide system of trauma care in Victoria: effect on patient survival. *Med J Aust* 2008;189(10):546-50.
23. Urquhart DM, Edwards ER, Graves SE, et al. Characterisation of orthopaedic trauma admitted to adult level 1 trauma centres. *Injury* 2006;37(2):120-7.
24. Wilson JTL, Pettigrew LEL, Teasdale GM. Structured Interviews for the Glasgow Outcome Scale and the Extended Glasgow Outcome Scale: Guidelines for Their Use. *J Neurotrauma* 1998;15(8):573-85.
25. Gabbe BJ, Simpson PM, Lyons RA, et al. How well do principal diagnosis classifications predict disability 12 months postinjury? *Inj Prev* 2015;21(e1):e120-6.
26. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095-108.
27. Van Beeck EF, Larsen CF, Lyons RA, et al. Guidelines for the conduction of follow-up studies measuring injury-related disability. *J Trauma* 2007;62(2):534-50.
28. Gabbe BJ, Simpson PM, Lyons RA, et al. Association between the number of injuries sustained and 12-month disability outcomes: evidence from the injury-VIBES study. *PloS one* 2014;9(12):e113467.
29. Kilgo PD, Osler TM, Meredith W. The worst injury predicts mortality outcome the best: rethinking the role of multiple injuries in trauma outcome scoring. *J Trauma* 2003;55(4):599-606; discussion 06-7.

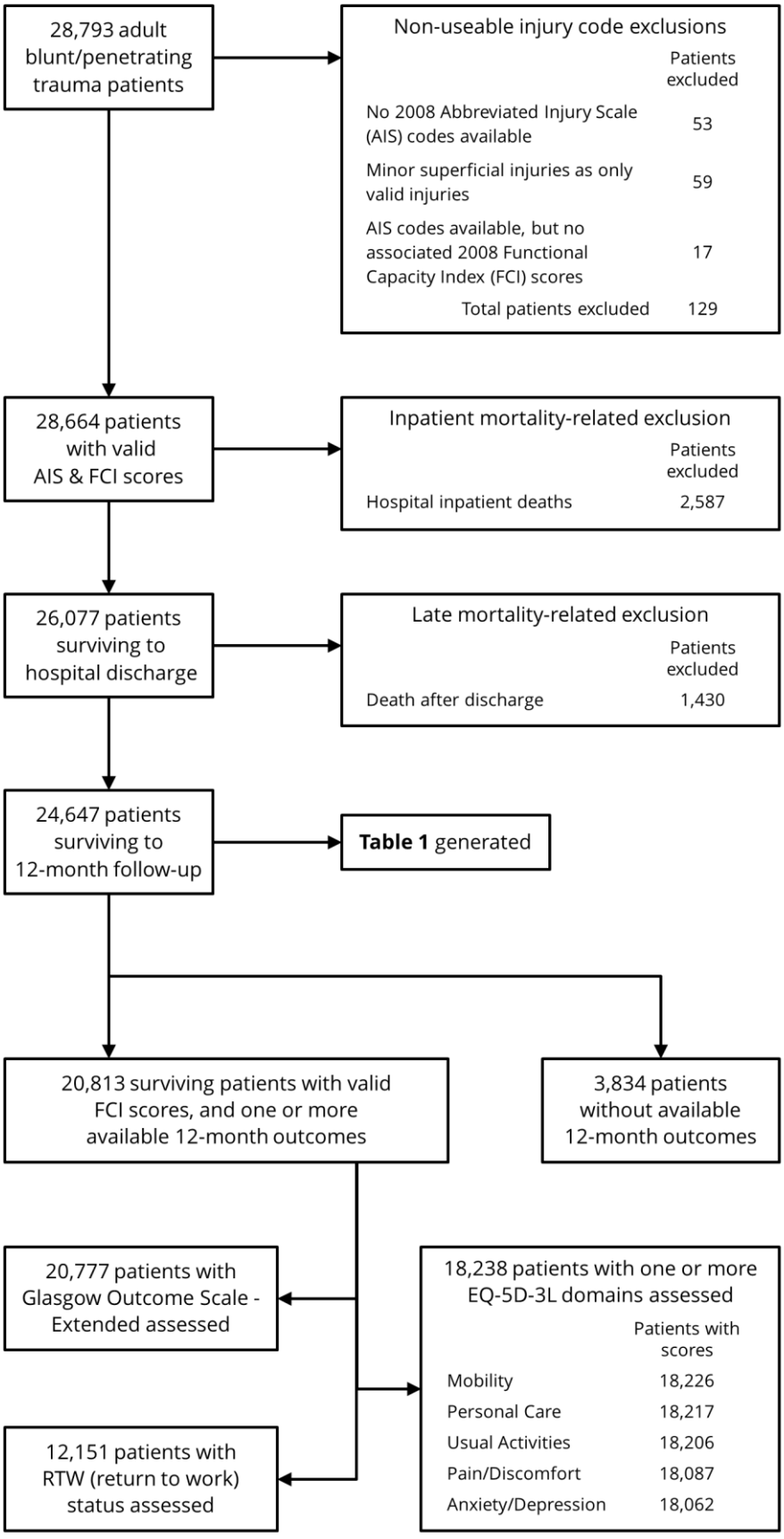
30. Baker SP, O'Neill B, Haddon W, et al. The Injury Severity Score: a method for describing patients with multiple injuries and evaluating trauma care. *J Trauma* 1974;14(3):187-96.
31. Osler T, Baker SP, Long W. A modification of the Injury Severity Score that both improves accuracy and simplifies scoring. *J Trauma* 1997;43(6):922-25.
32. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006;25:127-41.
33. Bertolini G, D'Amico R, Nardi D, et al. One model, several results: the paradox of the Hosmer-Lemeshow goodness-of-fit test for the logistic regression model. *J Epidemiol Biostatist* 2000;5(4):251-53.
34. Hosmer D, Hosmer T, Le Cessie F, et al. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16:965-80.
35. Gönen M. Analyzing Receiver Operating Characteristic Curves with SAS. Cary, NC: SAS Institute Inc. 2007.
36. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav* 2017;2(1):6-10.
37. Pink B. Socio-Economic Indexes for Areas (SEIFA). Canberra: Australian Bureau of Statistics, 2013.
38. Gabbe BJ, Simpson PM, Cameron PA, et al. Long-term health status and trajectories of seriously injured patients: A population-based longitudinal study. *PLoS Med* 2017;14(7):e1002322.
39. Barnes J, Thomas P. Quality of life outcomes in a hospitalized sample of road users involved in crashes. *Annu Proc Assoc Adv Automot Med* 2006;50:253-68.
40. Lyons RA, Kendrick D, Towner EM, et al. Measuring the population burden of injuries--implications for global and national estimates: a multi-centre prospective UK longitudinal study. *PLoS Med* 2011;8(12):e1001140.

FIGURE LEGENDS

Figure 1. Flow diagram showing derivation of the study dataset, including the number of patients with available data for each outcome measure used.

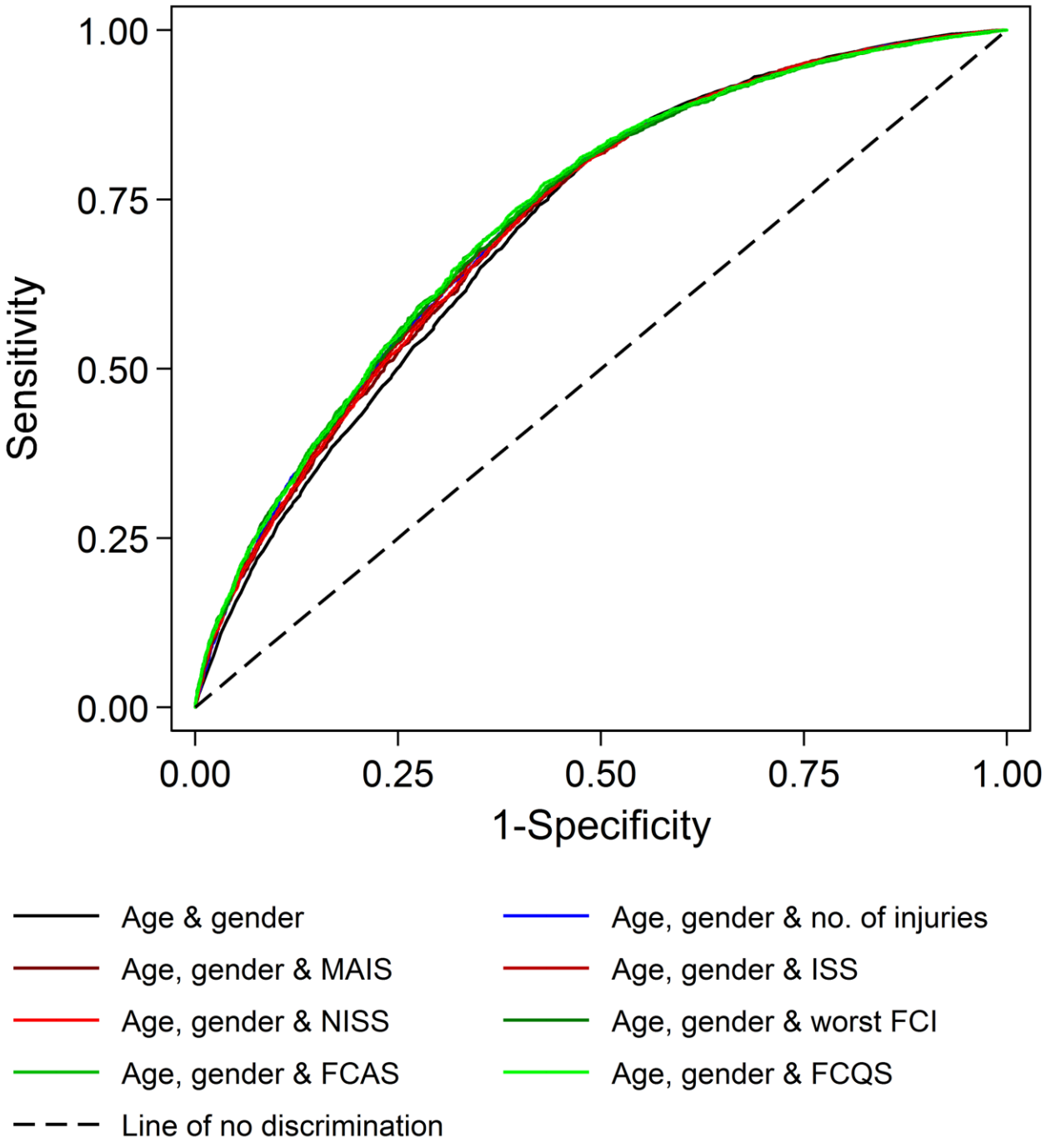
Figure 2. Receiver operating characteristic (ROC) curves for models used in predicting the EQ-5D personal care item in the training dataset (n=12,196 patients).

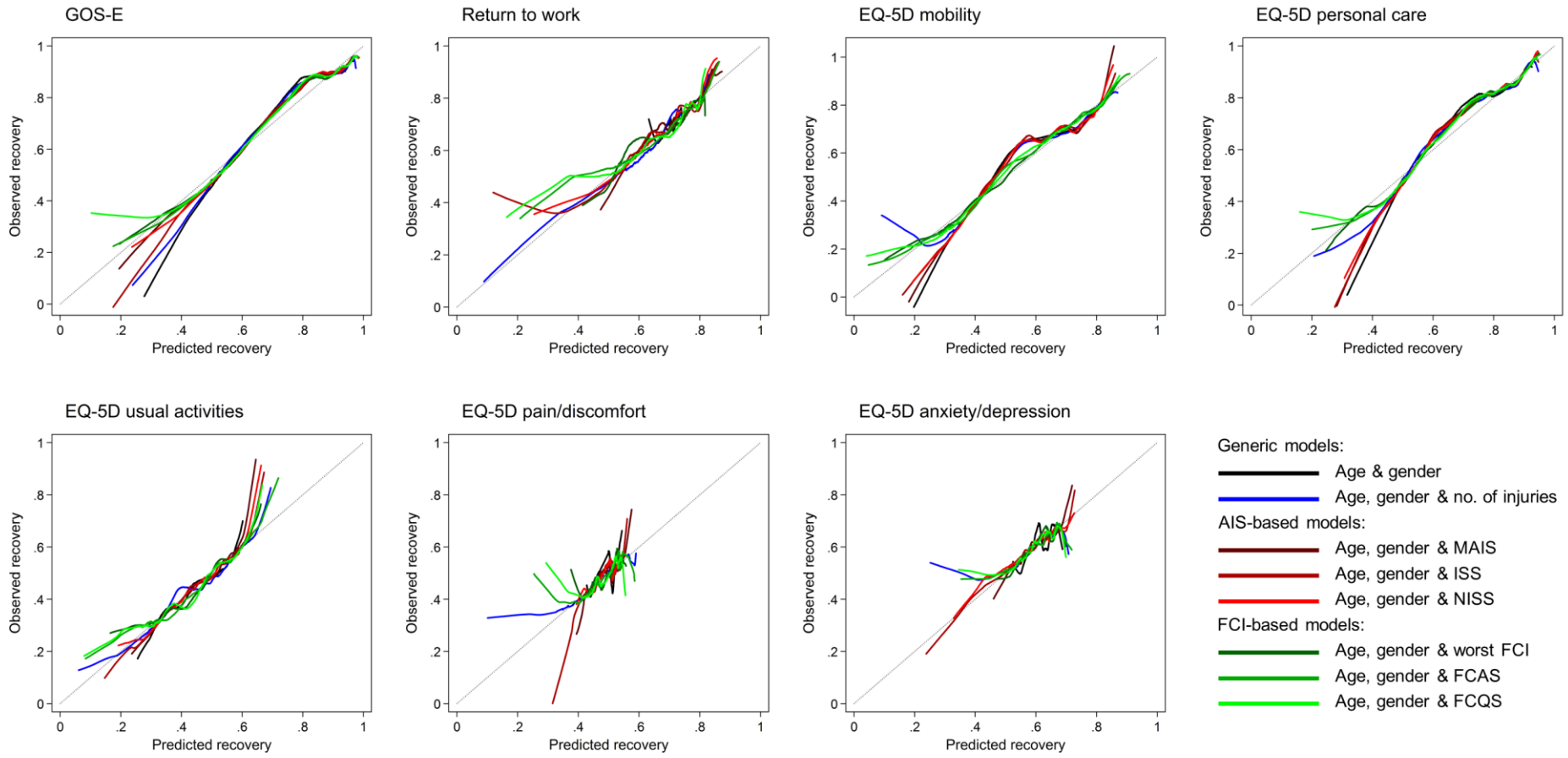
Figure 3. Plots illustrating predicted versus observed recovery in the testing datasets for each outcome variable evaluated. The 45° line shown in each sub-figure represents perfect fit of each model.



[a] AIS - 2008 Abbreviated Injury Scale

[b] FCI - 2008 predictive Functional Capacity Index


1.00
0.75
0.50
0.25
0.00
Sensitivity
0.00
0.25
0.50
0.75
1.00
1-Specificity
Age & gender
Age, gender & MAIS
Age, gender & NISS
Age, gender & FCAS
Line of no discrimination
Age, gender & no. of injuries
Age, gender & ISS
Age, gender & worst FCI
Age, gender & FCQS


GOS-E
Return to work
EQ-5D mobility
EQ-5D personal care
EQ-5D usual activities
EQ-5D pain/discomfort
EQ-5D anxiety/depression
Observed recovery
Predicted recovery
0
.2
.4
.6
.8
1
Generic models:
Age & gender
Age, gender & no. of injuries
AIS-based models:
Age, gender & MAIS
Age, gender & ISS
Age, gender & NISS
FCI-based models:
Age, gender & worst FCI
Age, gender & FCAS
Age, gender & FCQS