The Spine Journal 19 (2019) 703–710

Clinical Study

Highlighting discrepancies in walking prediction accuracy for patients with traumatic spinal cord injury: an evaluation of validated prediction models using a Canadian Multicenter Spinal Cord Injury Registry

Philippe Phan, MD, PhD[a,b,c,*], Brandon Budhram, BSc[a], Qiong Zhang, MSc[d,e], Carly S. Rivers, PhD[d], Vanessa K. Noonan, PhD[d,e], Tova Plashkes, PT[d], Eugene K. Wai, MD, MSc[a,b,c], Jérôme Paquet, MD[f], Darren M. Roffey, PhD[a,c], Eve Tsai, MD, PhD[a,c,g], Nader Fallah, PhD[d,e] The RHSCIR Network

[a] *Ottawa Combined Adult Spinal Surgery Program, The Ottawa Hospital, 1053 Carling Ave, Ottawa, ON K1Y 4E9, Canada*
[b] *Division of Orthopaedic Surgery, Department of Surgery, Faculty of Medicine, University of Ottawa, The Ottawa Hospital, 1053 Carling Ave, Ottawa, ON K1Y 4E9, Canada*
[c] *Clinical Epidemiology Program, The Ottawa Hospital,, 1053 Carling Ave, Ottawa, ON K1Y 4E9, Canada*
[d] *Rick Hansen Institute, Blusson Spinal Cord Centre, 6400-818 W. 10th Ave, Vancouver, BC V5Z 1M9, Canada*
[e] *The University of British Columbia, 2329 West Mall, Vancouver, BC V6T 1Z4, Canada*
[f] *Département Sciences Neurologiques, Pavillon Enfant-Jésus, CHU de Québec, 1401 18e rue, Québec, QC G1J 1Z4, Canada*
[g] *Division of Neurosurgery, Department of Surgery, Faculty of Medicine, University of Ottawa, The Ottawa Hospital, 1053 Carling Ave, Ottawa, ON K1Y 4E9, Canada*

Received 6 June 2018; revised 27 August 2018; accepted 27 August 2018

Abstract

BACKGROUND CONTEXT: Models for predicting recovery in traumatic spinal cord injury (tSCI) patients have been developed to optimize care. Several models predicting tSCI recovery have been previously validated, yet recent findings question their accuracy, particularly in patients whose prognoses are the least predictable.

PURPOSE: To compare independent ambulatory outcomes in AIS (ASIA [American Spinal Injury Association] Impairment Scale) A, B, C, and D patients, as well as in AIS B+C and AIS A+D patients by applying two existing logistic regression prediction models.

STUDY DESIGN: A prospective cohort study.

PARTICIPANT SAMPLE: Individuals with tSCI enrolled in the pan-Canadian Rick Hansen SCI Registry (RHSCIR) between 2004 and 2016 with complete neurologic examination and Functional Independence Measure (FIM) outcome data.

OUTCOME MEASURES: The FIM locomotor score was used to assess independent walking ability at 1-year follow-up.

METHODS: Two validated prediction models were evaluated for their ability to predict walking 1-year postinjury. Relative prognostic performance was compared with the area under the receiver operating curve (AUC).

RESULTS: In total, 675 tSCI patients were identified for analysis. In model 1, predictive accuracies for 675 AIS A, B, C, and D patients as measured by AUC were 0.730 (95% confidence interval [CI] 0.622–0.838), 0.691 (0.533–0.849), 0.850 (0.771–0.928), and 0.516 (0.320–0.711), respectively. In 160 AIS B+C patients, model 1 generated an AUC of 0.833 (95% CI 0.771–0.895), whereas model 2 generated an AUC of 0.821 (95% CI 0.754–0.887). The AUC for 515 AIS A+D

FDA device/drug status: Not applicable.

Author disclosures: ***PP:*** Nothing to disclose. ***BB:*** Nothing to disclose. ***QZ:*** Nothing to disclose. ***CR:*** Nothing to disclose. ***VKN:*** Nothing to disclose. ***TP:*** Nothing to disclose. ***EKW:*** Nothing to disclose. ***JP:*** Nothing to disclose. ***DMR:*** Nothing to disclose. ***ET:*** Nothing to disclose. ***NF:*** Nothing to disclose.

The disclosure key can be found on the Table of Contents and at www.TheSpineJournalOnline.com.

* Corresponding author. The Ottawa Hospital, Civic Campus, 1053 Carling Ave, Room J-155, Ottawa, ON K1Y 4E9, Canada. Tel.: (613) 761-5168; fax: (613) 761-4661.

E-mail address: pphan@toh.ca (P. Phan).

https://doi.org/10.1016/j.spinee.2018.08.016

patients was 0.954 (95% CI 0.933–0.975) with model 1 and 0.950 (0.928–0.971) with model 2. The difference in prediction accuracy between the AIS B+C cohort and the AIS A+D cohort was statistically significant using both models (p=.00034; p=.00038). The models were not statistically different in individual or subgroup analyses.

CONCLUSIONS: Previously tested prediction models demonstrated a lower predictive accuracy for AIS B+C than AIS A+D patients. These models were unable to effectively prognosticate AIS A+D patients separately; a failure that was masked when amalgamating the two patient populations. This suggests that former prediction models achieved strong prognostic accuracy by combining AIS classifications coupled with a disproportionately high proportion of AIS A+D patients.

Keywords: Functional outcomes; Logistic regression; Predictive accuracy; Prognosis; Walking; Traumatic spinal cord injury.

Introduction

The burden of traumatic spinal cord injury (tSCI) has been long recognized, both on individual and societal levels. These injuries often lead to irreversible motor, sensory, and autonomic damage, which may compromise functional ability, such as independent walking [1]. Following the initial traumatic and neurologic injury, the ability to predict long-term walking ability assists clinicians in providing patients and their families with realistic expectations regarding recovery, rehabilitation, and the extent of physical and/or financial burden [2,3].

The demand for effective prognostication in tSCI has lent itself to numerous studies evaluating recovery using prediction models. Van Middendorp et al. [4] used logistic regression (LR) models to accurately predict ambulatory outcomes from five clinical variables. This prediction model was initially tested on European data, and it was externally validated in the United States [5] and Australia [6]. Most recently, Hicks et al. [7] were able to validate this five-variable predictive rule using data from the Rick Hansen Spinal Cord Injury Registry (RHSCIR), a Canadian multicenter database [8]. Furthermore, they were able to generate an abridged three-variable model to predict walking ability with a similar level of accuracy based on an individual's Functional Independence Measure (FIM) score [9,10] at 1-year postinjury.

Despite the reported success of recent prognostic models, it is important to note that they were developed for all severities of tSCI as a collective whole, from the most severe injuries (AIS [American Spinal Injury Association] Impairment Scale A) to the least severe injuries (AIS D). Hicks et al. [7] observed that the sample populations in many of these studies consisted predominantly of AIS A+D patients, likely biasing their models' predictive abilities. Our objective was to examine the effect of such a bias on the prognostic accuracy of the model in those whose outcomes are least clinically predictable, patients with AIS B+C injuries, for whom outcome prediction is most useful. The importance of validating clinical prediction rules in the context of tSCI patients has been recently highlighted [11].

Our hypothesis is that the prognostication of AIS B+C patients is less accurate than that of AIS A+D, and that former models achieve high predictive accuracy by means of sampling bias. As such, we aim to examine and compare the predictive accuracy of the van Middendorp et al. [4] five-variable and Hicks et al. [7] three-variable prediction models (hereafter referred to as "van Middendorp model" and "Hicks model") in the individual AIS classifications, that is, AIS A, B, C, and D using the Hicks cohort. Furthermore, we will compare and contrast the predictive ability of these two models in the subgroups of AIS A+D patients as well as AIS B+C patients.

Methods

Analysis cohort

Participants for this study were obtained from the RHSCIR (2004–2016), which includes 18 acute care and 13 rehabilitation hospitals across Canada. The RHSCIR is a prospective observational registry created to answer research questions and to facilitate the implementation of best practices. Full details of the RHSCIR have been published elsewhere [8]. All participating sites obtained approval from their local research ethics board before enrolling participants. Any person who is the age of majority and receives treatment for a new tSCI at a participating RHSCIR site is eligible for inclusion in the registry. A core dataset is collected for all registrants, and a detailed dataset, including follow-up questionnaires performed at 1, 2, 5, and 10 years following injury is collected for those who provide written informed consent. Participants with complete neurologic and functional data as below formed the analysis cohort. This cohort is the same as used in Hicks et al. (see their paper for full cohort details).

Outcome measures

Neurologic severity (ie, AIS) and level of injury at time of admission were assessed within 15 days of injury according to the International Standards for Neurological

Classification of Spinal Cord Injury (ISNCSCI) [12]. Independent walking ability was assessed by the interview FIM at ≥12 months postinjury, by which time neurologic and functional outcomes are known to plateau [13−17].

Prediction modeling

We applied the same five predictive variables to the RHSCIR database as previously described by van Middendorp et al. [4] and Hicks et al.. These prognostic variables were age (dichotomized at 65 years old), motor scores of the quadriceps femoris (L3) and gastrocsoleus (S1) muscles, and light touch sensation of the corresponding L3 and S1 dermatomes. Using the cohort from Hicks et al. [7] (see their paper for full cohort details), we tested the van Middendorp and Hicks models in the individual AIS categories of A, B, C, and D patients, as well as in the subgroups AIS A+D and AIS B+C patient categories. For each model, classification accuracy, sensitivity, specificity, and the area under the receiver-operating curve (AUC, a measure of diagnostic performance via confidence intervals) were calculated. A p value of $<.05$ was considered statistically significant. All statistical analyses were performed using SPSS (version 23) and Rx64 (version 3.1).

Results

Demographic and injury details

The analysis cohort included 675 participants; of these, 210 (31.1%) were AIS A, 51 (7.6%) AIS B, 109 (16.1%) AIS C, and 305 (45.2%) AIS D. See Table 1 for full details. A comparison between the van Middendorp and Hicks model cohorts is reported elsewhere. Comparison of the subgroups used in this work and the original van Middendorp model cohort was not possible as data were not reported by these subgroups.

Prediction models

Prediction rule scores to determine probabilities of walking independently 1-year postinjury for both the van Middendorp and Hicks models were calculated individually for each of the AIS classifications (A, B, C, and D), as well as for the AIS A+D and AIS B+C subgroups. Concordance matrices used to assess the degree of agreement between two variables can be found in Table 2, with an accompanying density plot showing prediction of walking or not walking at 1-year postinjury for the individual AIS classes (Fig. 1).The scatter plot demonstrates a clear predictive bias for AIS A and AIS D, as visualized by an accumulation of not walking and walking patients, respectively.

Individual AIS classification prognostication

The overall classification accuracy, sensitivity, specificity, and AUCs generated by each of the tested models are summarized in Table 3 for AIS A, B, C, D, as well as for AIS A+D and AIS B+C subgroups.

Using the van Middendorp model, the AUCs for AIS A, B, C, and D were 0.730 (0.622−0.838), 0.691 (0.533−0.849), 0.850 (0.771−0.928), and 0.516 (0.320−0.711), respectively. Using the Hicks model, the AUC for AIS A, B, C, and D were 0.730 (0.621−0.839), 0.714 (0.565−0.863), 0.840 (0.747−0.933), and 0.519 (0.307−0.731), respectively. For reference, the overall classification accuracies for AIS A, B, C, and D using the van Middendorp

Table 1
Study and participant data for patients extracted from the RHSCIR for use in logistic regression models

	Patient cohort, AIS A and D (n=515)	Patient cohort, AIS B and C (n=160)
Setting	31 Canadian SCI centers	
Inclusion period	2004-2016	
Sex (male)	400 (78%)	126 (79%)
Age at time of injury, mean (SD)	47.2 (18.0)	45.1 (18.6)
Severity of neurological deficit at time of injury		
AIS A	210 (41%)	
AIS B		51 (32%)
AIS C		109 (68%)
AIS D	305 (59%)	
Lower extremity motor score, mean (SD)	26.4 (21.8)	8.0 (9.7)
Duration until first neurological evaluation	3.4 (4.0)	3.1 (3.9)
Mechanism of injury		
Assault, blunt	12 (2.3%)	1 (0.6%)
Assault, penetrating	12 (2.3%)	5 (3.1%)
Fall	219 (42.5%)	73 (45.6%)
Sports	91 (17.7%)	38 (23.8%)
Transport	149 (28.9%)	32 (20.0%)
Other traumatic	25 (4.9%)	9 (5.6%)
Surgical correction	82 (16.2%)	15 (9.6%)

AIS, American Spinal Injury Association Impairment Scale; SCI, spinal cord injury; SD, standard deviation.

Table 2
Concordance matrices for AIS A, B, C, and D individually, and AIS A+D and B+C cohorts, analysed by the van Middendorp and Hicks simplified logistic regression models

			Predicted (n)	
Class	Model	Actual outcome (n)	Cannot walk	Walk
AIS A (n=210)	van Middendorp	Cannot walk	183	2
		Walk	15	10
	Hicks	Cannot walk	183	2
		Walk	16	9
AIS B (n=51)	van Middendorp	Cannot walk	29	3
		Walk	11	8
	Hicks	Cannot walk	29	3
		Walk	11	8
AIS C (n=109)	van Middendorp	Cannot walk	14	14
		Walk	8	73
	Hicks	Cannot walk	16	12
		Walk	8	73
AIS D (n=305)	van Middendorp	Cannot walk	0	8
		Walk	0	297
	Hicks	Cannot walk	0	8
		Walk	0	297
AIS A+D (n=515)	van Middendorp	Cannot walk	181	12
		Walk	16	306
	Hicks	Cannot walk	181	12
		Walk	25	297
AIS B+C (n=160)	van Middendorp	Cannot walk	35	25
		Walk	17	83
	Hicks	Cannot walk	37	23
		Walk	18	82

AIS, American Spinal Injury Association Impairment Scale.

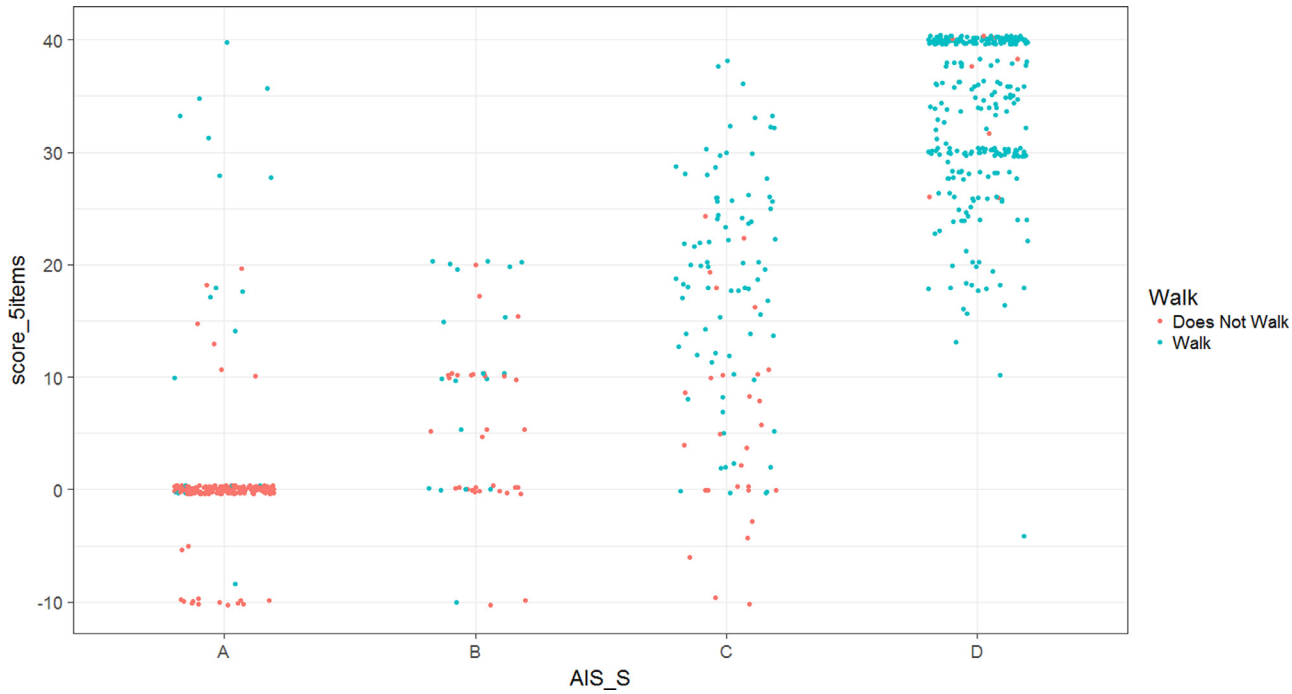

Fig. 1. Scatter plot depicting walking vs. not walking ability for individual AIS classifications (A, B, C, and D)

Table 3
Predictive accuracy of the van Middendorp and Hicks logistic regression models for individual AIS classifications and AIS A+D and B+C subgroups

AIS class	Model	OCA (%)	Sensitivity (%)	Specificity (%)	AUC (95% CI)
AIS A (n=210)	van Middendorp	91.49	37.23	98.55	0.730 (0.622–0.838)
	Hicks	91.10	33.80	98.57	0.730 (0.621–0.839)
AIS B (n=51)	van Middendorp	71.52	37.13	90.29	0.691 (0.533–0.849)
	Hicks	72.70	38.22	90.84	0.714 (0.565–0.863)
AIS C (n=109)	van Middendorp	78.52	88.62	48.24	0.850 (0.771–0.928)
	Hicks	80.51	89.84	52.90	0.840 (0.747–0.933)
AIS D (n=305)	van Middendorp	97.38	100	0	0.516 (0.320–0.711)
	Hicks	97.38	100	0	0.519 (0.307–0.731)
A+D (n=515)	van Middendorp	94.56	95.02	93.80	0.954 (0.933–0.975)
	Hicks	92.68	92.04	93.78	0.950 (0.928–0.971)
B+C (n=160)	van Middendorp	73.70	82.89	58.64	0.833 (0.771–0.895)
	Hicks	74.28	81.99	61.73	0.821 (0.754–0.887)

OCA, overall classification accuracy; AUC, area under the receiver operating characteristics curve; CI, confidence interval.

model were 91.49%, 71.52%, 78.52%, and 97.38%, respectively.

AIS A+D and AIS B+C subgroup prognostication

For AIS A+D, the AUCs for the van Middendorp and Hicks models were 0.954 (95% confidence interval [CI] 0.933–0.975) and 0.950 (95% CI 0.928–0.971), respectively. For AIS B+C, the AUCs for the van Middendorp and Hicks models were 0.833 (95% CI 0.771–0.895) and 0.821 (95% CI 0.754–0.887), respectively.

Model comparisons

The difference in AUC between AIS A+D and AIS B+C cohorts was statistically significant using both the van Middendorp and Hicks models (p=.00038). When comparing between the two models, the difference of AUCs was not statistically significant for AIS A+D (p=.131) or AIS B+C (p=.448). The receiver operating characteristic (ROC) curves for both models comparing AIS A+D with AIS B+C are shown in Fig. 2.

Discussion

In this study of the RHSCIR, we investigated the utility of two multivariable prediction models in a specific cohort of AIS B+C patients recovering from tSCI. The multinational validated, five-variable model proposed by van Middendorp et al. [4] poorly predicted recovery in AIS A and AIS D patients individually, but was of higher predicative accuracy when combining AIS A+D patients into a single cohort. When comparing individuals with AIS A+D with those with AIS B+C, there was a considerably higher degree of prognostication present with the AIS A+D cohort, in keeping with our hypothesis. These results were similarly reflected utilizing the simplified, three-variable model from Hicks et al. [7], collectively suggesting that current prediction models have achieved erroneously high predictive accuracies by combining all tSCI patients with all severities of injury into a single cohort.

Part of this bias is related to the limitation of ROC to be used as a diagnostic test. Very high sensitivity and specificity and consequently high AUC values can be artificially obtained by simply, increasing sample size in diagonal values in the concordance matrix. Fig. 3 presents a hypothetical example assuming a different distribution of AIS (panel A) where the sample has a majority of AIS A+D patients (and less AIS B+C patients); therefore, both sensitivity and specificity are very high. However, if the same sample size has a higher proportion of AIS B+C, that is, those whose outcomes are more variable and for whom a prognostic model is most useful, the sensitivity and specificity are lower, demonstrating how easily bias may affect results (panel B).

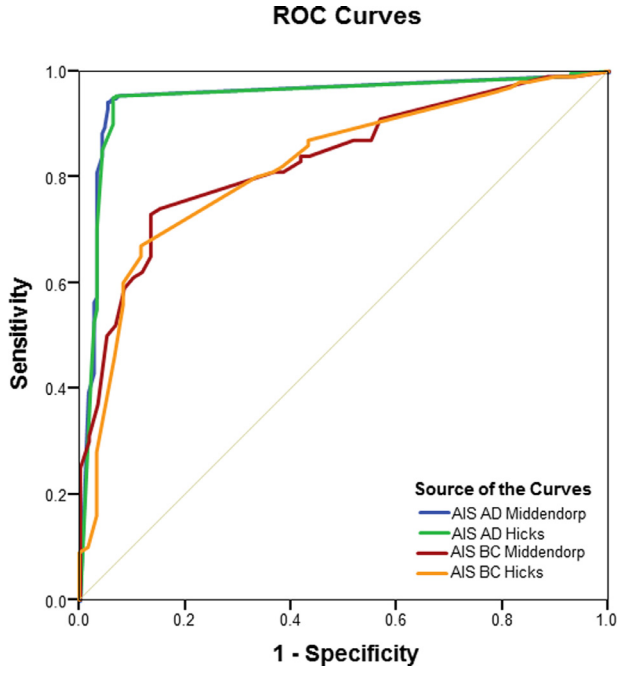

Fig. 2. Area under the operating curve for AIS A+D and B+C cohorts, using both the van Middendorp and Hicks models.

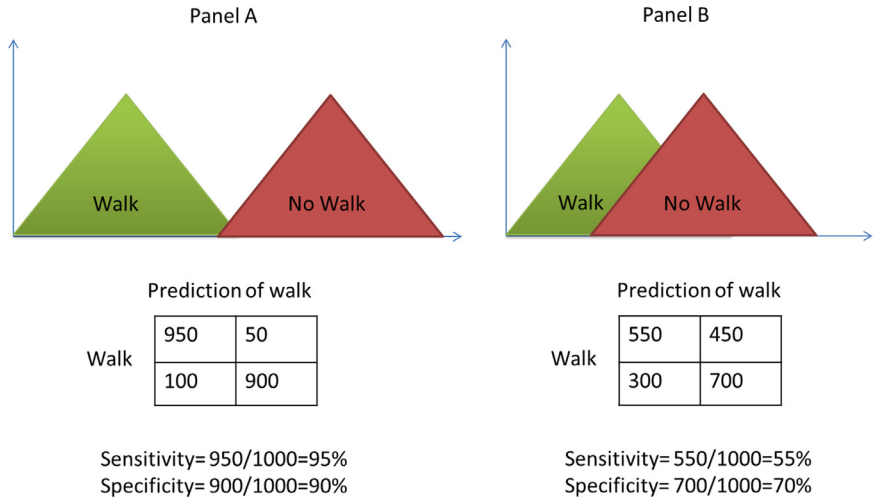

Fig. 3. Effect of distribution of severity of injury on the performance of the predictive model.

Applying currently established prediction models, we tested our hypothesis that the prognostic capability of tSCI prediction rules for walking 1-year postinjury may be compromised by the preponderant and predictable nature of AIS A (for whom walking is unlikely) and AIS D (for whom walking is likely) patients. Of recent prediction rules, the van Middendorp five-variable prediction model has garnered considerable international attention, and its prognostic capability has been evaluated and validated in numerous studies [4–7]. This is clinically relevant, as external validation and proper methodology are crucial in filtering the recently increasing number of clinical prediction rules [11]. However, although past studies have shown high predictive accuracy prognosticating tSCI patients collectively, our subgroup analyses confirm that there is a clear predilection for well-defined outcomes in AIS A+D patients with regard to ambulatory outcomes, as demonstrated by an AUC of 0.954 using the van Middendorp et al. model [4]. This is in contrast with AIS B+C patients, where the AUC using the same model was 0.833. Similarly, Hicks et al. [7] demonstrated an AUC of 0.950 for AIS A+D and 0.821 for AIS B+C. These findings are in accordance with the suggestion by Hicks et al. [7] that there exists a discrepancy in the prediction of ambulatory outcomes for each AIS class by logistic regression models. This discrepancy might be understated by the misleadingly high prediction accuracy of those models due to the high prevalence of patients with predictable outcomes (AIS A+D) in national registries.

In addition to the differences in subgroup analysis, there was a notable failure of both models to prognosticate patients within individual AIS classifications. This is most prominent in reference to the individual AIS A and AIS D populations. Both models demonstrated poor sensitivity (eg, 37.2% in the van Middendorp model) and relatively low AUC for AIS A, effectively underestimating the number of patients that would walk 1-year postinjury. Although it makes intuitive sense that AIS A patients (ie, patients with complete SCI) would experience unfavorable ambulatory outcomes, the ability of some individuals to experience significant functional outcomes is supported in the literature [18]. Conversely, for AIS D, both models predicted that all 305 patients would walk 1-year postinjury (sensitivity 100%, specificity 0%), showing a clear propensity for false negatives in this population. Using the van Middendorp model, the poor AUCs for AIS A and AIS D individually (0.730 and 0.516, respectively) are not mirrored in their overall classification accuracies (91.49% and 97.38%, respectively), presenting a dichotomy between the two measurements. Furthermore, these AUCs are effectively masked when combining these two groups into a single A+D cohort (AUC of 0.954).These findings suggest that although AIS A and AIS D patients have naturally predictable outcomes, when assessed individually they are poorly prognosticated by current models.

Although AUC is a common tool to measure predictive accuracy in a binary classification system [19], its reported values in former studies describing tSCI prognostication have been misleading. With this in mind, recent literature has suggested that AUC may actually be a misrepresentative measure of predictive distribution model performance, and that its values require compounding with other measures (eg, sensitivity, specificity) in order to truly delineate prediction outcomes [20]. Given that recent clinical prediction rules have relied primarily on AUC to determine prognostic ability, the addition of other predictive measures may provide supplemental benefit in defining accurate ambulatory outcomes. Other methods to validate predictive models that are being developed, such as a prediction-recall curve, may be more appropriate in future evaluations [21].

Despite the proposal by Hicks et al. [7] that a three-variable prediction model could provide nearly equivalent predictive accuracy to the validated five-variable model, a notion that was echoed in the recent commentary by Nater

and Fehlings [11], our analyses suggest that perhaps generating prediction models with increased variable count and/or using non-LR methodologies could improve tSCI prognostication. Additional variables that could be of interest include spinal MRI findings and serum biomarkers. Wilson et al. [22] were able to devise a regression prediction models for long-term ambulatory and functional outcomes utilizing intramedullary signal characteristics on spinal MRI as a variable. The salience of spinal MRIs in these patients has been supported in numerous studies since [23–25]. With respect to serum biomarkers, Pouw et al. [26] noted post-tSCI variations in the following proteins: S-100β, neuron specific enolase, phosphorylated neurofilament (heavy subunit), and glial fibrillary acidic protein. These biomarkers, among numerous others, have been consistently described in the literature as important factors in tSCI prognostication [27–29]. With respect to methodology, the limitation to binary analysis with LR has been described [30]. As such, more complex methods have been applied to predict functional outcomes in the tSCI population. Using neural network analysis, a form of artificial intelligence, Belliveau et al. [31] created a model to effectively predict ambulatory independence and self-care at 1-year postinjury, rivaling the predictive accuracy of LR. The growing body of evidence supporting the use of additional variables and/or different methodologies is crucial in the development of future prognostic models.

Strengths of our study include the scope of centers covered by the RHSCIR, providing a good representation of the true tSCI patient population on a national level. Furthermore, this is the first paper highlighting clinically critical shortcomings in validated prediction models [4–7] by demonstrating not only that the success of these models lies in the skewed proportion of A+D patients (comprising the majority of the sample population) but also that these models are ineffective at prognosticating patients when evaluating patients within a single AIS classification level.

This study has some limitations. Using a binary outcome (not walking vs walking) to simulate independent ambulatory functions at a fixed time point (1-year postinjury) as opposed to a continuous outcome (eg, ambulatory improvements over an extended time period) may not be truly indicative of patient recovery. Further, the van Middendorp model analyzed outcome using the Spinal Cord Independence Measure, whereas the Hicks model used the FIM. However, dichotomizing the outcome of not walking versus walking reduces the subtleties in difference between the two measures. Additionally, this study has a relatively limited sample size, owing to the difficulty of extracting complete patient data such as neurologic sensory scores from a multicenter database. Further, AIS B+C are the least common tSCI, and as a result, led to a small sample size. Last, given that the RHSCIR is subject to response rates and participant willingness, our sample population may not be entirely representative of the true patient population; the results should be validated in other SCI datasets.

Conclusions

This study demonstrates the limited accuracy of validated models for prognostication of walking in AIS B+C patients in contrast to AIS A+D patients. It also underlines the inability of these models to prognosticate patients within a given AIS classification (eg, AIS A and AIS D individually). This highlights a possible misperception from former publications that prediction models have high and equivalent prediction accuracies for all AIS classes. In light of these findings, it may be beneficial for future prediction models to incorporate additional relevant variables or use more advanced analytical methodologies including artificial intelligence.

Acknowledgments

Data for this research were collected using the Global Research Platform (http://www.rhigrp.net/).

Neurology data for this research was cleaned using the RHI ISNCSCI Algorithm, a free tool (http://www.isncscialgorithm.com/).

We would like to thank the RHSCIR network and all the participating local RHSCIR sites: GF Strong Rehabilitation Centre, Vancouver General Hospital, Foothills Hospital, Glenrose Rehabilitation Hospital, Royal Alexandra Hospital, University of Alberta Hospital, Royal University Hospital, Saskatoon City Hospital, Winnipeg Health Sciences Centre, Toronto Western Hospital, Toronto Rehabilitation Institute, St. Michael's Hospital, Sunnybrook Health Sciences Centre, Hamilton General Hospital, Hamilton Health Sciences Regional Rehabilitation Centre, Victoria Hospital (London), University Hospital (London), Parkwood Hospital (London), The Ottawa Hospital Rehabilitation Centre, The Ottawa Hospital Civic Campus, Hôpital de l'Enfant Jésus, Institut de réadaptation en deficience physique de Quebec, Centre de réadaptation Lucie-Bruneau, Institut de réadaptation Gingras-Lindsay-de-Montréal, Hôpital du Sacre Cœur de Montréal, Nova Scotia Rehabilitation Centre, QEII Health Sciences Centre, Saint John Regional Hospital, Stan Cassidy Centre for Rehabilitation, St. John's Health Sciences Centre, and L.A. Miller Rehabilitation Centre. This study and the RHSCIR are supported by funding from Health Canada, Western Economic Diversification Canada, and the Governments of Alberta, British Columbia, Manitoba, and Ontario. The funding agencies had no role in the study or study reporting.

References

[1] Noonan VK, Fingas M, Farry A, Baxter D, Singh A, Fehlings MG, et al. Incidence and prevalence of spinal cord injury in Canada: a national perspective. Neuroepidemiology 2012;38:219–26. doi: 10.1159/000336014.

[2] Kirshblum SC, Priebe MM, Ho CH, Scelza WM, Chiodo AE, Wuermser L-A. Spinal cord injury medicine. 3. Rehabilitation phase after acute spinal cord injury. Arch Phys Med Rehabil 2007;88(Suppl 3):S62–70. doi: 10.1016/j.apmr.2006.12.003.

[3] Krueger H, Noonan VK, Trenaman LM, Joshi P, Rivers CS. The economic burden of traumatic spinal cord injury in Canada. Chronic Dis Inj Can 2013;33:113–22.

[4] van Middendorp JJ, Hosman AJ, Donders AR, et al. A clinical prediction rule for ambulation outcomes after traumatic spinal cord injury: a longitudinal cohort study. Lancet 2011;377:1004–10. doi: S0140-6736(10)62276-3 [pii]\r10.1016/S0140-6736(10)62276-3.

[5] Malla R. External validation study of a clinical prediction rule for ambulation outcomes after traumatic spinal cord injury. 2013.

[6] Van Silfhout L, Peters A, Graco M, Schembri R, Nunn A, Berlowitz D. Validation of the Dutch clinical prediction rule for ambulation outcomes in an inpatient setting following traumatic spinal cord injury. Spinal Cord 2015;54:614–8. doi: 10.1038/sc.2015.201.

[7] Hicks KE, Zhao Y, Fallah N, et al. A simplified clinical prediction rule for prognosticating independent walking after spinal cord injury: a prospective study from a Canadian multicenter spinal cord injury registry. Spine J 2017;17:1383–92. doi: 10.1016/j.spinee.2017.05.031.

[8] Noonan VK, Kwon BK, Soril L, et al. The Rick Hansen Spinal Cord Injury Registry (RHSCIR): a national patient-registry. Spinal Cord 2012;50:22–7. doi: 10.1038/sc.2011.109.

[9] UB Foundation Activities, Inc; FIM® instrument (FIM), UB Foundation Activities, Inc.: Amherst, NY; 2012.

[10] Keith RA, Granger C V, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for rehabilitation. Adv Clin Rehabil 1987;1:6–18.

[11] Nater A, Fehlings MG. Clinical prediction rules: the importance of the validation phase. Spine J 2017;17:1393–6. doi: 10.1016/j.spinee.2017.06.002.

[12] Kirshblum SC, Waring W, Biering-Sorensen F, et al. Reference for the 2011 revision of the international standards for neurological classification of spinal cord injury. J Spinal Cord Med 2011;34:547–54. doi: 10.1179/107902611X13186000420242.

[13] Ditunno JF. The John Stanley Coulter Lecture. Predicting recovery after spinal cord injury: a rehabilitation imperative. Arch Phys Med Rehabil 1999;80(4):361–4.

[14] Ditunno JF, Cohen ME, Hauck WW, Jackson AB, Sipski ML. Recovery of upper-extremity strength in complete and incomplete tetraplegia: a multicenter study. Arch Phys Med Rehabil 2000;81:389–93. doi: 10.1053/mr.2000.3779.

[15] Cheung WY, Arvinte D, Wong YW, Luk KDK, Cheung KMC. Neurological recovery after surgical decompression in patients with cervical spondylotic myelopathy – a prospective study. Int Orthop 2008;32:273–8. doi: 10.1007/s00264-006-0315-4.

[16] Burns A, Yee J, Flett HM, Guy K, Cournoyea N. Impact of benchmarking and clinical decision making tools on rehabilitation length of stay following spinal cord injury. Spinal Cord 2013;51:165–9. doi: 10.1038/sc.2012.91.

[17] Pretz CR, Kozlowski AJ, Charlifue S, Chen Y, Heinemann AW. Using rasch motor FIM individual growth curves to inform clinical decisions for persons with paraplegia. Spinal Cord 2014;52:671–6. doi: 10.1038/sc.2014.94.

[18] Bracken MB, Shepard MJ, Collins WF, et al. A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal-cord injury. Results of the Second National Acute Spinal Cord Injury Study. N Engl J Med 1990;322:1405–11. doi: 10.1056/NEJM199005173222001.

[19] Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Casp J Intern Med Spring 2013;4:627–35. doi: 10.1017/CBO9781107415324.004.

[20] Lobo JM, Jiménez-valverde A, Real R. AUC: a misleading measure of the performance of predictive distribution models. Glob Ecol Biogeogr 2008;17:145–51. doi: 10.1111/j.1466-8238.2007.00358.x.

[21] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 2015:e0118432. doi: 10.1371/journal.pone.0118432.

[22] Wilson JR, Grossman RG, Frankowski RF, et al. A clinical prediction model for long-term functional outcome after traumatic spinal cord injury based on acute clinical and imaging factors. J Neurotrauma 2012;29:2263–71. doi: 10.1089/neu.2012.2417.

[23] Chandra J, Sheerin F, Lopez De Heredia L, et al. MRI in acute and subacute post-traumatic spinal cord injury: pictorial review. Spinal Cord 2012 Jan;50:2–7. doi: 10.1038/sc.2011.107.

[24] Fehlings MG, Martin AR, Tetreault LA, et al. A clinical practice guideline for the management of patients with acute spinal cord injury: recommendations on the role of baseline magnetic resonance imaging in clinical decision making and outcome prediction. Glob Spine J 2017;7:221S–30S. doi: 10.1177/2192568217703089.

[25] Ahuja CS, Wilson JR, Nori S, et al. Traumatic spinal cord injury. Nat Rev Dis Prim 2017 Apr 27;3:17018. doi: 10.1038/nrdp.2017.18.

[26] Pouw MH, Hosman AJF, van Middendorp JJ, Verbeek MM, Vos PE, van de Meent H. Biomarkers in spinal cord injury. Spinal Cord 2009 Jul;47:519–25. doi: 10.1038/sc.2008.176.

[27] Kwon BK, Stammers AMT, Belanger LM, et al. Cerebrospinal fluid inflammatory cytokines and biomarkers of injury severity. J Neurotrauma 2010;27:669–82.

[28] Loy DN, Kim JH, Xie M, Schmidt RE, Trinkaus K, Song S-K. Diffusion tensor imaging predicts hyperacute spinal cord injury severity. J Neurotrauma 2007;24:979–90. doi: 10.1089/neu.2006.0253.

[29] Kwon BK, Streijger F, Fallah N, et al. Cerebrospinal fluid biomarkers to stratify injury severity and predict outcome in human traumatic spinal cord injury. J Neurotrauma 2017;34:567–80. doi: 10.1089/neu.2016.4435.

[30] Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. Lancet 1995 Oct 21;346:1075–9. doi: 10.1016/S0140-6736(95)91746-2.

[31] Belliveau T, Jette AM, Seetharama S, et al. Developing artificial neural network models to predict functioning one year after traumatic spinal cord injury. Arch Phys Med Rehabil 2016;97:1663–8 e3. doi: 10.1016/j.apmr.2016.04.014.