ORIGINAL ARTICLE

Reliability of the spine adverse events severity system (SAVES) for individuals with traumatic spinal cord injury

RA Glennie[1], VK Noonan[1,2], N Fallah[1,2], SE Park[1,2], NP Thorogood[2], A Cheung[2], CG Fisher[1], MF Dvorak[1,2] and JT Street[1]

Study design: Test–retest analysis.

Objectives: To determine the intra- and inter-rater reliability of the Spine Adverse Events Severity System for Spinal Cord Injury (SAVES-SCI) in patients with traumatic SCI.

Setting: Quaternary care spine program in Vancouver, Canada.

Methods: Ten hypothetical patient cases were developed. The cases were completed by 10 raters (seven physicians, one nurse, one physiotherapist and one researcher) who were asked to identify and grade the severity of adverse events using SAVES-SCI twice with 1-week interval. Intra- and inter-rater reliability were calculated using kappa statistics and intraclass correlation coefficients (ICC).

Results: Intra-rater reliability for both identifying and grading AEs were high with all AEs (kappa greater than 0.6) except for bone implant, diathermy burn, massive blood loss, myocardial infarction, neurological deterioration, pressure ulcer, return to operating room and tracheostomy requirment. The inter-rater reliability measured with ICC were all above 0.6 for identifying and grading intraoperative, pre and postoperative AEs and consequences of SCI.

Conclusions: The SAVES-SCI demonstrated acceptable intra-and inter-rater reliability for a majority of the AEs. Further clarification and definition of some of the AEs as well as provision of sample training cases for clincians would assist in reducing measurement errors. The SAVES-SCI is a useful tool to assess and capture AEs in patients with acute traumatic SCI.

Sponsorship: Funded by Rick Hansen Institute and Health Canada.

INTRODUCTION

Acute adverse events (AE) are common in patients with traumatic spinal cord injury (tSCI). Their reported incidence varies depending on the methods used to identify and report them. Retrospective reviews of tSCI patient records have reported incidences ranging from 50–60%[1–3] whereas the Spine Adverse Events Severity System (SAVES), a prospective data collection instrument, found that almost 80% of patients with tSCI experienced at least one AE during acute care.[4] Accurate AE reporting is essential for the development of clinical care guidelines, appropriate allocation of resources and meaningful clinical and multicenter research collaboration. The disparity in published incidence reporting highlights a critical need for a reliable method to detect and report AEs in acute care.

The SAVES was developed as a spine-specific instrument to prospectively identify, categorize and classify AEs in elective spinal surgery for degenerative conditions.[5] The development, content validation and reliability testing of the SAVES instrument in the general spine surgery population are outlined elsewhere.[5] The SAVES, however, does not reflect the unique challenges of patients with tSCI, and thus may not sufficiently capture relevant AEs specific to this population. For instance, surgery for individuals with tSCI may be associated with higher incidences of tracheostomy, neurological

deterioration and wound complications than the degenerative spinal population. Patients with tSCI also frequently encounter AEs as a consequence of the injury itself such as pressure ulcers, spasticity, autonomic dysreflexia and neuropathic pain. An instrument that collects meaningful AE data in tSCI patients must incorporate these and other conditions that are specific to this population.

At our institution, a group of clinicians and researchers obtained permission to modify the SAVES and included AEs regularly observed among patients with tSCI. The modified instrument, termed the SAVES-SCI, was developed following an analysis of the original tool[6] in patients with tSCI[4] and a review of the literature on acute AEs commonly experienced by patients with tSCI. Clinical experts reviewed the initial list of AEs and added AEs considered clinically important to tSCI. See Supplementary Appendix 1 for the SAVES-SCI instrument. The purpose of this study was to determine the intra- and inter-rater reliability of the SAVES-SCI in a tSCI population among physicians, nurses, physiotherapists and research staff.

METHODS

Development of patient cases

Ten unique patient cases were developed by an occupational therapist and a spine surgeon, neither of whom were involved in the reliability testing. The cases were hypothetical clinical scenarios involving patients with tSCI.

[1]Division of Spine, Department of Orthopedics, University of British Columbia, Vancouver, British Columbia, Canada and [2]Rick Hansen Institute, Vancouver, British Columbia, Canada
Correspondence: Dr JT Street, Division of Spine, Department of Orthopedics, Blusson Spinal Cord Centre, University of British Columbia, Room 6180, 818 West 10th Avenue, 2733 Heather Street, Vancouver, British Columbia, Canada V5Z 1M9.
E-mail: John.Street@vch.ca

We based the scenarios on actual patients admitted in the previous 12 months to reflect characteristics representative of our tSCI population. For example, as 80% of new tSCI cases in British Columbia have been male,[7] 8 of the 10 clinical scenarios described male patients. The scenarios were presented in a patient chart format and depicted the patient's demographic, injury and medical background information, type of surgical and acute care received and an overall clinical summary of their experience throughout the hospital stay, including AEs. See Supplementary Appendix 2 for a sample patient case.

Data collection

The SAVES-SCI consists of three AE categories: intraoperative, pre- or post-surgical intervention, consequences secondary to SCI and others. The grading of AEs was simplified to dichotomize the response. A score of '0' was given to AEs deemed not to impact patient outcome and length of stay, and a score of '1' to those AEs deemed to have such an impact.

Seven fellowship-trained spine surgeons and three clinicians/researchers (one nurse, one physiotherapist and one researcher), all with at least 15 years of clinical or research experience, were invited to participate in the reliability study. Each rater was instructed to identify and grade all AEs in each of the 10 patient cases twice with a 1-week interval in between to provide information on intra-rater reliability.

Data analysis

For the assessment of intra-rater reliability, a simple kappa was calculated. The value of 0.6 was used as a benchmark for high reliability.[8] The strength of agreement given by kappa values was interpreted as follows according to the guideline by Landis and Koch: $\leqslant 0$, poor agreement; 0–0.2, slight agreement; 0.2–0.4, fair agreement; 0.4–0.6, moderate agreement; 0.6–0.8, substantial agreement; 0.8–1.0, almost perfect agreement.[8]

For the assessment of inter-rater reliability, each rater's responses for the 10 patient cases were assessed to determine if the AEs were identified and graded correctly compared with an answer key. A panel of six senior spine surgeons with experience using the SAVES had previously examined the cases and created the answer key. Inter-rater reliability was determined by calculating two types of multi-rater kappa. Fleiss' extension of the kappa, also called the generalized kappa, is used to measure agreement between three or more raters.[9] Conger's kappa is an exact kappa, and is another way of measuring agreement between three or more raters.[10] A higher kappa suggests low variability between responses and high reliability of consistently identifying or grading the particular AE correctly among different raters. As with intra-rater reliability, a value of 0.6 was used as a cut-off to indicate high reliability.[8] As Conger kappa values were almost identical to the Fleiss values, only Fleiss kappa values are reported in this study. For pre- and post-surgical intervention AEs and consequences of SCI, several AEs were differentiated by the date they occurred and treated as separate events for the rating and analysis. However, for data reporting, these distinctions were collapsed and the median kappa value was reported for each type of AEs.

Inter-rater reliability was assessed again using a two-way random effect of the intraclass correlation coefficient (ICC).[11] For this analysis, we added all of the raters' data together using their total AE score to compare their first evaluation with their second evaluation. A high ICC would indicate that the SAVES-SCI can be reliably used over time among different raters. Analyses were performed using R version 2.15 (R Development Core Team, Vienna, Austria) and SPSS version 21.0 (SPSS, Chicago, IL, USA). This study was approved by the ethics board at the study university and hospital.

RESULTS

There were ten raters in total who completed the SAVES-SCI for the ten hypothetical patient cases. The complete list of AEs with corresponding kappa and *P*-values is presented in Supplementary Appendix 3. We could not assess the intra- or inter-rater agreement for grading some AEs as some raters did not provide a grade, such that caused the response had no variance.

Intraoperative AEs

There were 10 intraoperative AEs represented in the hypothetical patient cases. See Table 1 for the intra- and inter-rater reliability of identifying the presence of these AEs and Table 2 for the reliability of grading their severity. Intra-rater agreement for identifying AEs was generally high, ranging from $\kappa = 0.65$–1.00 whereas inter-rater agreement was lower, ranging from $\kappa = 0.12$–0.92. Kappa values for grading AEs was lower than that of identifying AEs for both intra-rater (range 0.44–1.00) and inter-rater reliability (range 0.00–0.82). Allergic reaction was the only AE that had almost perfect intra- and inter-rater reliability for both its identification and grading. Inter-rater reliability measured using ICC was higher for identifying (ICC = 0.94, 95% CI = 0.92–0.96) than grading intraoperative AEs (ICC = 0.79, 95% CI = 0.69–0.86), which was consistent with the results of the kappa analysis.

Pre- or post-surgical intervention AEs

There were 19 unique pre- or post-surgical intervention AEs included in the patient cases. Only 8 of the 10 raters identified and graded AEs in this group. See Table 3 for the intra- and inter-rater reliability of identifying the presence of these AEs and Table 4 for the reliability of grading their severity. Intra-rater agreement for identifying AEs was moderate to high, ranging from $\kappa = 0.49$–1.00. Like intraoperative AEs, inter-rater agreement was lower than intra-rater AEs with a range of $\kappa = 0.00$–0.93. When kappa statistics were compared between grading and identifying AEs, intra-rater agreement for grading (range 0.36–1.00) was similar to that of identifying AEs but generally lower (range 0.00–0.87) for inter-rater agreement. No AEs in this group had consistently poor or almost perfect intra- and inter-rater agreement for both identification and grading. Unlike the kappa statistics, ICC was higher for grading (ICC = 0.70, 95% CI = 0.53–0.81) than identifying AEs (ICC = 0.61, 95% CI = 0.39–0.75). Compared with intraoperative AEs, raters identified and graded pre- or post-surgical intervention AEs less consistently with each other as indicated by the lower ICC.

Table 1 Strength of intra-/inter-rater reliability for identifying the presence of intraoperative AEs

Strength of agreement	Intra-rater reliability (100 observations)	Inter-rater reliability (20 observations)
Almost perfect (0.80< $\kappa \leqslant$1.00)	Allergic reaction Cord injury Dural tear Hardware malposition Massive blood loss Fall	Allergic reaction Dural tear Hardware malposition
Substantial (0.60< $\kappa \leqslant$0.80)	Bone implant Vascular injury Airway/ventilation Diathermy burn	Cord injury Fall Diathermy burn
Moderate (0.40< $\kappa \leqslant$0.60)	None	Massive blood loss
Fair (0.20< $\kappa \leqslant$0.40)	None	Airway/ventilation
Slight (0.00< $\kappa \leqslant$0.20)	None	Bone implant Vascular injury
Poor (κ<0.00)	None	None
Not available	None	None

Abbreviation: AE, adverse event.

Consequences secondary to SCI

There were five AEs in the consequences secondary to SCI category. Only eight of the ten raters completed this section. See Table 5 for the intra-and inter-rater reliability of identifying the presence of these AEs

Table 2 Strength of intra-/inter-rater reliability for identifying and grading the severity of intraoperative AEs

Strength of agreement	Intra-rater reliability (100 observations)	Inter-rater reliability (20 observations)
Almost perfect (0.80 < κ ⩽ 1.00)	Allergic reaction Cord injury Fall	Allergic reaction Dural tear
Substantial (0.60 < κ ⩽ 0.80)	Dural tear Airway/ventilation	Cord injury Fall
Moderate (0.40 < κ ⩽ 0.60)	Bone implant Hardware malposition Massive blood loss Diathermy burn	None
Fair (0.20 < κ ⩽ 0.40)	None	Hardware malposition Massive blood loss Airway/ventilation Diathermy burn
Slight (0.00 < κ ⩽ 0.20)	None	Bone implant
Poor (κ < 0.00)	None	Vascular injury
Not available	Vascular injury	None

Abbreviation: AE, adverse event.

and Table 6 for the reliability of grading their severity. All of these AEs were identified with almost perfect intra-rater agreement (range 0.84–0.93). Inter-rater agreement was also quite high (range 0.60–0.93), but lower than intra-rater agreement as for intraoperative and pre- or post-surgical intervention AEs. Intra-rater agreements for grading AEs in this group were substantial to almost perfect (range 0.66–0.92). The inter-rater agreement was lower compared to identifying these AEs (range 0.31–0.87). Mood disturbance and renal calculi had almost perfect intra- and inter-rater agreement for both its identification and grading. Similar to kappa statistics, ICC was also lower for grading (ICC = 0.79, 95% CI = 0.69–0.86) than identifying AEs (ICC = 0.94, 95% CI = 0.92–0.96). Compared to pre- or post-surgical intervention AEs, raters identified and graded consequences secondary to SCI more consistently with each other themselves as indicated by the higher ICC. ICC was almost the same for identifying consequences secondary to SCI and identifying intraoperative AEs, but raters more accurately graded consequences of SCI than intraoperative AEs.

DISCUSSION

This study examined the intra- and inter-rater reliability of identifying and grading the severity of intraoperative, pre- and post-surgical intervention AEs and consequences of SCI using the SAVES-SCI. The majority of these AEs were reliably assessed with κ ⩾ 0.6, both for the same rater over time (intra-rater reliability) and between raters (inter-rater reliability). In comparison to published AE measures for the SCI population which are tailored towards community settings,[12,13]

Table 3 Strength of intra-/inter-rater reliability for identifying the presence of pre-/post-surgical intervention AEs

Strength of agreement	Intra-rater reliability (80 observations)	Inter-rater reliability (20 observations)
Almost perfect (0.80 < κ ⩽ 1.00)	Deep vein thrombosis Delirium Dysphagia Humerus fracture Pneumonia Respiratory distress Urinary tract infection	Delirium Dysphagia Pneumonia Urinary tract infection
Substantial (0.60 < κ ⩽ 0.80)	Cardiac arrest/failure/arrhythmia Construct failure without loss of correction Myocardial infarction Post-operative neuropathic pain due to hardware malposition Tracheostomy required	Deep Vein thrombosis Myocardial infarction
Moderate (0.40 < κ ⩽ 0.60)	Neurological deterioration ⩾ 1 motor grade in ASIA Motor Scale Pressure ulcer Return to operating room for change of screw/hardware malposition	Post-operative neuropathic pain due to hardware malposition
Fair (0.20 < κ ⩽ 0.40)	None	Cardiac arrest/failure/arrhythmia Humerus fracutre Neurological deterioration ⩾ 1 motor grade in ASIA Motor Scale Pressure ulcer Respiratory distress
Slight (0.00 < κ ⩽ 0.20)	None	Construct Failure without loss of correction Hematoma Tracheostomy required Return to operating room for change of screw/hardware malposition
Poor (κ < 0.00)	None	Multiorgan/system failure
Not Available	Fever of Unknown origin Hematoma Massive blood loss Multiorgan/system failure	Fever of unknown origin Massive blood loss

Abbreviation: AE, adverse event.

Table 4 Strength of intra-/inter-rater reliability for grading the severity of pre-/post-surgical intervention AEs

Strength of agreement	Intra-rater reliability (80 observations)	Inter-rater reliability (20 observations)
Almost Perfect ($0.80 < \kappa \leqslant 1.00$)	Delirium	None
	Humerus fracture	
	Respiratory Distress	
Substantial ($0.60 < \kappa \leqslant 0.80$)	Cardiac arrest/failure/arrhythmia	Deep Vein thrombosis
	Construct failure without loss of correction	Delirium
	Deep vein thrombosis	Pneumonia
	Dysphagia	
	Pneumonia	
	Post-operative neuropathic pain due to hardware malposition	
	Pressure ulcers	
	Urinary tract infection	
Moderate ($0.40 < \kappa \leqslant 0.60$)	Neurological deterioration $\geqslant 1$ motor grade in ASIA Motor Scale	Dysphagia
	Tracheostomy required	Myocardial infarction
		Postoperative neuropathic pain due to hardware malposition
		Urinary tract infection
Fair ($0.20 < \kappa \leqslant 0.40$)	Myocardial infarction	Humerus fracture
		Neurological deterioration $\geqslant 1$ motor grade in ASIA Motor Scale
		Respiratory distress
Slight ($0.00 < \kappa \leqslant 0.20$)	None	Cardiac arrest/failure/arrhythmia
		Construct failure without loss of correction
		Pressure ulcers
		Tracheostomy required
Poor ($\kappa < 0.00$)	None	Hematoma
		Return to operating room for change of screw/hardware malposition
		Multiorgan/system failure
Not available	Fever of unknown origin	Fever of unknown origin
	Hematoma	Massive blood loss
	Return to operating room for change of screw/hardware malposition	
	Massive blood loss	
	Multiorgan/system failure	

Abbreviation: AE, adverse event.

Table 5 Strength of intra-/inter-rater reliability for identifying the presence of consequences secondary to SCI

Strength of agreement	Intra-rater reliability (80 observations)	Inter-rater reliability (20 observations)
Almost perfect ($0.80 < \kappa \leqslant 1.00$)	Autonomic dysreflexia	Autonomic dysreflexia
	Joint contractures	Joint contractures
	Mood disturbance requiring treatment	Mood disturbance requiring treatment
	Renal calculi	Renal calculi
	Ventilatory compromise	
Substantial ($0.60 < \kappa \leqslant 0.80$)	None	None
Moderate ($0.40 < \kappa \leqslant 0.60$)	None	Ventilatory compromise
Fair ($0.20 < \kappa \leqslant 0.40$)	None	None
Slight ($0.00 < \kappa \leqslant 0.20$)	None	None
Poor ($\kappa < 0.00$)	None	None
Not available	None	None

Abbreviation: SCI, spinal cord injury.

the SAVES-SCI is better suited to measure AEs in an acute setting as it includes a section on intraoperative AEs.

Intra- versus Inter-rater reliability

In all three groups of AEs, intra-rater reliability was higher than the corresponding inter-rater reliability, which was expected. This is consistent with other reliability studies of tools measuring acute AE data.[14,15] Only 9 out of 34 AEs had kappa less than 0.6; these were bone implant, diathermy burn, hardware malposition, massive blood loss, myocardial infarction, neurological deterioration, pressure ulcer, return to operating room and tracheostomy requirement. Variability among raters may have resulted from differences in raters' clinical experience or expertise working with AEs in SCI, as our raters came from a range of disciplines. Sharek et al.,[15] reported a greater ability to identify AEs with increased exposure to the charts. Future studies should compare intra- and inter-rater reliability within the same discipline as disciplines vary in their familiarity with the various AEs, as well as the way their impact is graded.

Table 6 Strength of intra-/inter-rater reliability for identifying and grading the severity of consequences secondary to SCI

Strength of agreement	Intra-rater reliability (80 observations)	Inter-rater reliability (20 observations)
Almost perfect ($0.80 < \kappa \leqslant 1.00$)	Mood disturbance requiring treatment	Mood disturbance requiring treatment
	Renal calculi	Renal calculi
	Ventilatory compromise	
Substantial ($0.60 < \kappa \leqslant 0.80$)	Autonomic dysreflexia	None
	Joint contractures	
Moderate ($0.40 < \kappa \leqslant 0.60$)	None	Autonomic dysreflexia
Fair ($0.20 < \kappa \leqslant 0.40$)	None	Joint contractures
		Ventilatory compromise
Slight ($0.00 < \kappa \leqslant 0.20$)	None	None
Poor ($\kappa < 0.00$)	None	None
Not available	None	None

Abbreviation: SCI, spinal cord injury.

It would also be worth exploring whether a specific discipline should be responsible for certain sections of the SAVES-SCI. Surgeons would be ideal candidates to complete the intraoperative AE section whereas nurses or therapists may be better suited to monitor and record pre- and post-surgical intervention AEs depending on workflow. This sharing of responsibilities would lessen the burden of time required to complete the form, a frequently quoted barrier in implementation of best practices,[16,17] and would facilitate the use of the SAVES-SCI in a busy acute care setting.

Identifying versus grading AEs

In general, we found that the intra- and inter-rater reliability of grading AEs were inferior to identifying them. This suggests that there is ambiguity in assessing the severity of AEs, which could result from the variation among raters with their clinical experience. The literature reports mixed results in this area. Sharek et al.,[15] reported higher kappa for identifying than grading AEs, whereas Klopotowska et al.,[14] reported higher agreement for grading than identifying AEs for both intra-rater and inter-rater reliability. The results from the latter study may differ from ours as Klopotowska et al.,[14] had a senior physician–pharmacist team as reviewers whereas our study involved reviewers from various disciplines.[14] Also, Klopotowska et al.,[14] limited their patient population to those over the age of 65 who were receiving more than five medications, and restricted the type of AEs collected to those that arose from medications whereas our study used no such limits.[14] Individuals using the SAVES-SCI should be trained on how to grade AEs, particularly for the pre- and post-surgical intervention AEs with low kappa values. This can facilitate better capturing the impact of AEs on length of stay and would be important when examining the costs of acute care for the tSCI population.

Limitations

There are limitations to consider when interpreting the results of this study. This study examined the reliability of the SAVES-SCI, but further research is needed to examine the validity of this new tool. In addition, the analyses only included the raters that identified the AEs and therefore, did not capture when a rater missed an AE. This occurs in the pre- or post-surgical intervention AE group where only 8 of the 10 raters identified AEs. It is important to look further into this issue since the inability to identify AEs suggests that specific training may be necessary.

It is also important to note that this study focuses on AEs common in the tSCI population. Rouleau et al.,[18] reported a difference between non-traumatic and tSCI patients in both their demographic characteristics and the AEs they experience, so future studies should examine the tool's reliability in non-traumatic SCI populations as well. Also, as AEs in the acute and rehabilitation settings are different, the SAVES-SCI should be tested and modified for a rehabilitation setting, if necessary.

Future directions

The presence of AEs increases acute length of stay and delays patient participation in rehabilitation.[19] A firm understanding of the incidence and severity of AEs in acute SCI is important to ensure that appropriate resources are allocated to minimize the incidence and impact on patient outcomes. The SAVES-SCI is being used to prospectively collect AE data on tSCI patients admitted to the study hospital and has demonstrated a superior ability to capture AE data than administrative ICD-10 codes.[20] Future studies examining the effect of raters' discipline (clinicians versus researchers), the validity of the SAVES-SCI and its use in differing patient populations (for example, non-traumatic SCI) and other settings (for example, rehabilitation) should all be considered.

CONCLUSION

The SAVES-SCI demonstrated acceptable reliability for the majority of its AEs. With further clarification of the definition of AEs and targeted training in identifying and grading AEs with low kappas, this tool could be implemented in acute clinical settings working with a tSCI population and could assist the clinical staffs to better identify and manage AEs.

DATA ARCHIVING

There were no data to deposit.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Fletcher DJ, Taddonio RF, Byrne DW, Wexler LM, Cayten CG, Nealon SM et al. Incidence of acute care complications in vertebral column fracture patients with and without spinal cord injury. Spine (Phila Pa 1976) 1995; **20**: 1136–1146.
2 Aito S, Gruppo Italiano Studio Epidemiologico Mielolesioni GISEM Group. Complications during the acute phase of traumatic spinal cord lesions. Spinal Cord 2003; **41**: 629–635.

3 Krassioukov AV, Furlan JC, Fehlings MG. Medical co-morbidities, secondary complications, and mortality in elderly with acute spinal cord injury. *J Neurotrauma* 2003; **20**: 391–399.

4 Street JT, Noonan VK, Cheung A, Fisher CG, Dvorak MF. Incidence of acute care adverse events and long-term health related quality of life in patients with traumatic spinal cord injury. *Spine J* 2013; **51**: 472–476.

5 Rampersaud YR, Neary MA, White K. Spine adverse events severity system: content validation and interobserver reliability assessment. *Spine (Phila Pa 1976)* 2010; **35**: 790–795.

6 Rampersaud YR, Moro ER, Neary MA, White K, Lewis SJ, Massicotte EM *et al.* Intraoperative adverse events and related postoperative complications in spine surgery: implications for enhancing patient safety founded on evidence-based protocols. *Spine (Phila Pa 1976)* 2006; **31**: 1503–1510.

7 Lenehan B, Street J, Kwon BK, Noonan V, Zhang H. The epidemiology of traumatic spinal cord injury in British Columbia, Canada. *Spine (Phila Pa 1976)* 2012; **37**: 321–329.

8 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.

9 Fleiss JL. Measuring nominal scale agreement among many rater. *Psychol Bull* 1971; **76**: 378–382.

10 Conger AJ. Integration and generalization of Kappas for multiple raters. *Psychol Bull* 1980; **88**: 322–328.

11 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–428.

12 Noreau L, Cobb J, Belanger LM, Dvorak MF, Leblond J *et al.* Development and assessment of a community follow-up questionnaire for the Rick Hansen spinal cord injury registry. *Arch Phys Med Rehabil* 2013; **94**: 1753–1765.

13 Kalpakjian CZ, Scelza WM, Forchheimer MB, Toussaint LL. Preliminary reliability and validity of a Spinal Cord Injury Secondary Conditions Scale. *J Spinal Cord Med* 2007; **30**: 131–139.

14 Klopotowska JE, Wierenga PC, Stuijt CC, Arisz L, Dijkgraaf MG, Kuks PF *et al.* Adverse drug events in older hospitalized patients: results and reliability of a comprehensive and structured identification strategy. *PLoS ONE* 2013; **8**: e71045.

15 Sharek PJ, Parry G, Goldmann D, Bones K, Hackbarth A, Resar R *et al.* Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Serv Res* 2011; **46**: 654–678.

16 Guihan M, Bosshart HT, Nelson A. Lessons learned in implementing SCI clinical practice guidelines. *SCI Nurs* 2004; **21**: 136–142.

17 Bloemen-Vrencken JH, de Witte LP, Engels JP, van den Heuvel WJ, Post MW. Transmural care in the rehabilitation sector: implementation experiences with a transmural care model for people with spinal cord injury. *Int J Integr Care* 2005; **5**: e02.

18 Rouleau P, Guertin PA. Traumatic and non-traumatic spinal cord-injured patients in Quebec, Canada. Part 2: biochemical profile. *Spinal Cord* 2010; **48**: 819–824.

19 Santos A, Gurling J, Dvorak MF, Noonan VK, Fehlings MG *et al.* Modeling the patient journey from injury to community reintegration for persons with acute traumatic spinal cord injury in a Canadian centre. *PLoS ONE* 2013; **8**: e72552.

20 Street JT, Thorogood NP, Cheung A, Noonan VK, Chen J, Fisher CG *et al.* Use of the Spine Adverse Events Severity System (SAVES) in patients with traumatic spinal cord injury. A comparison with institutional ICD-10 coding for the identification of acute care adverse events. *Spinal Cord* 2013; **51**: 472–476.

Supplementary Information accompanies this paper on the Spinal Cord website (http://www.nature.com/sc)