Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2016;97:1663-8

ORIGINAL RESEARCH

Developing Artificial Neural Network Models to Predict Functioning One Year After Traumatic Spinal Cord Injury

Timothy Belliveau, PhD,[a] Alan M. Jette, PT, PhD,[b] Subramani Seetharama, MD,[c] Jeffrey Axt, MBA,[d] David Rosenblum, MD,[e] Daniel Larose, PhD,[f] Bethlyn Houlihan, MSW, MPH,[b] Mary Slavin, PT, PhD,[b] Chantal Larose, PhD[g]

From the [a]Psychology Department, Hospital for Special Care, New Britain, CT; [b]Department of Health Policy and Management, Boston University School of Public Health, Boston, MA; [c]Physical Medicine and Rehabilitation Department, Hartford Hospital, Hartford, CT; [d]Information Technology Department, Hospital for Special Care, New Britain, CT; [e]Physical Medicine and Rehabilitation Department, Gaylord Hospital, Wallingford, CT; [f]Department of Mathematical Sciences, Central Connecticut State University, New Britain, CT; and [g]School of Business, State University of New York at New Paltz, New Paltz, NY.

Abstract

Objective: To develop mathematical models for predicting level of independence with specific functional outcomes 1 year after discharge from inpatient rehabilitation for spinal cord injury.

Design: Statistical analyses using artificial neural networks and logistic regression.

Setting: Retrospective analysis of data from the national, multicenter Spinal Cord Injury Model Systems (SCIMS) Database.

Participants: Subjects (N=3142; mean age, 41.5y) with traumatic spinal cord injury who contributed data for the National SCIMS Database longitudinal outcomes studies.

Interventions: Not applicable.

Main Outcome Measures: Self-reported ambulation ability and FIM-derived indices of level of assistance required for self-care activities (ie, bed-chair transfers, bladder and bowel management, eating, toileting).

Results: Models for predicting ambulation status were highly accurate (>85% case classification accuracy; areas under the receiver operating characteristic curve between .86 and .90). Models for predicting nonambulation outcomes were moderately accurate (76%−86% case classification accuracy; areas under the receiver operating characteristic curve between .70 and .82). The performance of models generated by artificial neural networks closely paralleled the performance of models analyzed using logistic regression constrained by the same independent variables.

Conclusions: After further prospective validation, such predictive models may allow clinicians to use data available at the time of admission to inpatient spinal cord injury rehabilitation to accurately predict longer-term ambulation status, and whether individual patients are likely to perform various self-care activities with or without assistance from another person.

Archives of Physical Medicine and Rehabilitation 2016;97:1663-8

After acute medical stabilization, individuals with spinal cord injury (SCI) are often confronted with a high level of uncertainty about their prospects for recovery of functional abilities. The early phase of inpatient rehabilitation represents an important opportunity for reevaluation of neurologic status and counseling about prognosis.[1] Many patients ask whether they will recover various functional abilities[2,3] (eg, walking ability, bowel and bladder control, sexual functioning) or whether they will need long-term assistance in the home setting. Previous research has consistently identified key motor and sensory examination findings, and the associated level of neurologic injury, that provide the most

Supported by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) (grant no. H133N10019). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this article do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.

Disclosures: none.

0003-9993/16/$36 - see front matter © 2016 by the American Congress of Rehabilitation Medicine
http://dx.doi.org/10.1016/j.apmr.2016.04.014

relevant neurologic information about prognosis for functional recovery.[4-11] Although estimates vary somewhat across studies, generalized projected outcomes have been determined for ambulation and other functional activities based largely on the patient's level of injury or the American Spinal Injury Association (ASIA) Impairment Scale (AIS)[12] classification obtained between 72 hours and 1 week postinjury, or both.[13-16]

In addition to prognostic information conveyed by single predictors such as AIS classification, clinical decision support tools[17,18] have emerged that aggregate information across predictors to improve the accuracy in projecting an individual patient's most likely functional outcomes. The use of such mathematical models and information technology to support clinical decision-making in rehabilitation medicine has been referred to as rehabilitation informatics, a field that "links the clinical expertise of health service providers with the technical expertise of information scientists to build new and innovative information technology applications to meet the needs of persons with neurologic disabilities."[18(p163)] A high level of predictive accuracy is possible when logistic regression models optimally weigh a combination of early predictors of a specific functional outcome, as shown by studies that developed mathematical models of subjects' projected walking ability at 6 months to 1 year postinjury.[19,20] For example, van Middendorp et al[20] showed that a simple, elegant logistic regression model with 5 variables (age dichotomized at 65y, muscle strength and light touch sensation at L3 and S1) assessed within 15 days of injury was highly accurate in the prediction of ambulation ability at 1 year postinjury. Similarly, Wilson et al[21] created a logistic regression model that was highly accurate in predicting long-term functional independence based on clinical and imaging data available within 3 days of injury. Using individual growth curve analyses, Pretz et al[22] generated algorithms that were robustly accurate in predicting patients' trajectories of motor recovery as measured by the motor subscale of the FIM.[23] One previous study[24] demonstrated greater than 80% sensitivity, specificity, and accuracy through the use of artificial neural networks (ANNs)[25] to predict ambulation status at the time of discharge from inpatient SCI rehabilitation.

The current study compares the accuracy and performance of 2 methods of predictive analytics (logistic regression and ANNs) to estimate longer-term ambulation and nonambulation outcomes, using outcome indices that are most directly relevant to questions about the probability of independent functioning versus the necessity of assistance from another person.

Methods

Participants

The cohort consisted of 3211 individuals who contributed data to the National Spinal Cord Injury Model Systems (SCIMS) Database per local institutional review board—approved procedures, between 2010 and 2014, who were at least 18 years of age at the time of traumatic SCI, for whom all ASIA neurologic exam motor scores were obtained at the time of rehabilitation admission, and who (per National Spinal Cord Injury Model Systems Database protocol) responded to questions about their functional status when they were 1 year posthospital discharge. To decrease heterogeneity in the duration of time between injury and the AIS examination done on admission to a SCIMS rehabilitation program, subjects were excluded from analyses if the length of time between their date of injury and admission to rehabilitation exceeded 9 weeks. These selection criteria yielded a dataset with 3142 subjects whose data were used for predictive analytics. AIS classifications were available for a subset of 3097 subjects, composed of 80% men and 20% women. Their ages ranged from 18 to 90 years (mean age ± SD, 41.5±17.1y). The mean duration ± SD between date of injury and admission to rehabilitation was 16.8±12.4 days. At the time of admission to rehabilitation, 41.2% of subjects were classified as AIS grade A, 13.1% as grade B, 18.9% as grade C, and 26.8% as grade D.

List of abbreviations:

AIS	**American Spinal Injury Association Impairment Scale**
ANN	**artificial neural network**
ASIA	**American Spinal Injury Association**
AUC	**area under the curve**
NLR	**negative likelihood ratio**
PLR	**positive likelihood ratio**
SCI	**spinal cord injury**
SCIMS	**Spinal Cord Injury Model Systems**

Variables

Input (independent) variables included subjects' age at the time of injury, sex, and ASIA exam manual motor testing scores. Output (dependent) variables consisted of subjects' responses to questions about their ambulation status, and their level of independence with transfers between the bed and chair, bowel management, bladder management, eating, and toileting. Per the National Spinal Cord Injury Model Systems Database outcome assessment protocol, in the context of a longer interview, all subjects were asked whether they were able to walk, with our without mobility aids, for 150ft in their home, 1 street block outside, and 1 flight of stairs. The remaining output (dependent) variables were derived from subjects' responses to FIM questions about bed-chair transfers, bladder and bowel management, eating, and toileting. Subjects' self-reported ability to perform each outcome activity was operationally defined as independent based on an FIM score of 7 (indicating the ability to do the activity without assistance from another person) or 6 (indicating modified independence involving the use of adaptive equipment but no help from another person). FIM scores of 1 through 5 (indicating that help from another person is required) were categorized as dependent.

Predictive models

A brief explanation of machine learning using ANNs[25-31] is available in supplemental appendix S1 (available online only at http://www.archives-pmr.org/). For each model, data from a randomly selected subset of 80% of the cases were used during the "network training" phase of analysis. In order to monitor for possible overfitting of the model to the training dataset, the accuracy and generalizability of each model were established on a cross-validation dataset composed of the remaining 20% of cases. Sensitivity analyses were then completed by generating a second set of ANNs using a random selection of 50% of the original cases and applying those networks to the total set of cases. To facilitate network training, input values for age and motor functioning were recoded as binary variables. Consistent with previous research,[20] subjects' ages at the time of SCI were recoded as a binary variable with values dichotomized at ≥65 years of age. The maximal score for every myotome (ie, the highest of the right- or left-sided motor score) was included. Motor strength scores were recoded from standard ASIA examination ratings from 0 to 5 into a binary variable with values dichotomized at $\geq$2, indicating that the person has, at minimum, the ability to perform active movement with gravity eliminated. The bilateral lower extremity (L2, L3, L4, L5, S1) motor score and the bilateral upper extremity (C5, C6, C7, C8, T1) motor score used for some of the predictive models were recoded from a 0- to 50-point scale to a 5-point scale (ranging from 1 to 5).

ANNs were generated using commercially available software.[a] Logistic regression analyses were completed using an IBM-SPSS Modeler,[b] with selection of independent variables in the regression models constrained by or essentially yoked to the same input variables used for the corresponding ANN analyses. The accuracy and performance of these predictive models were determined by percentage of correctly classified subjects (classification accuracy), true positive rate (sensitivity), true negative rate (specificity), proportion of positively classified subjects who were true positives (positive predictive value), and proportion of negatively classified subjects who were true negatives (negative predictive value). Likelihood ratios were also calculated. The positive likelihood ratio (PLR) is the proportion of subjects who achieved the functional outcome and were positively classified to the proportion of cases who did not achieve the functional outcome who were positively classified. The negative likelihood ratio (NLR) is the proportion of subjects who achieved the functional outcome who were negatively classified to the proportion of subjects who did not achieve the functional outcome who were negatively classified.[32] Another measure used to evaluate the performance of these predictive models is the area under the receiver operating characteristic curve (AUC), which compares sensitivity versus specificity across a range of values for the ability to predict a dichotomous outcome.[30,31,33]

Results

Descriptive analyses

Preliminary analysis consisted of a description of the number of subjects with each AIS classification and the percentage of these subjects whose responses were in the range of modified independence or independence for each activity at 1 year after hospital discharge. As would be expected, differential outcomes for these AIS categorical classifications were statistically significant (all Pearson χ^2 analyses, $P<.001$). However, as seen in table 1, estimation of future functional status based solely on AIS classification at the time of admission to rehabilitation leaves the potential for prognostic error, particularly for individuals whose examination results yield a classification of AIS grade B or C, whose functional outcomes are more variable.

Predictive models

The performance of ANN and logistic regression models for predicting these functional outcomes 1 year after hospital discharge is summarized in table 2 by the percentage of cases accurately classified and associated likelihood ratios, confidence intervals, and comparison of their respective AUCs. The specific input variables, prevalence of each outcome, performance characteristics of each predictive model, and sensitivity analyses are provided in supplemental tables S1 through S8 (see supplemental appendix S1).

ANN and logistic regression algorithms derived from subjects' age and the presence of at least gravity-eliminated motor strength at L2, L3, and S1 converged in their ability to accurately classify about 85% to 88% of subjects' self-reported ambulation status at 1 year after hospital discharge. This level of incremental improvement in predictive accuracy (ANN PLR=8.59, NLR=.17 for walking 150ft) occurs in a sample with a 39% to 42% base rate of self-reported ambulation activities. The models for predicting self-reported ability to ambulate achieved a somewhat higher negative predictive value than positive predictive value. The models for predicting nonambulation self-care activities had essentially moderate predictive power, with the best performance observed for the prediction of at least modified independence with bed-chair transfers (PLRs approximately 2.6, NLRs approximately .15). Performance of the ANN models remained stable when applied to a dataset that was held back for cross-validation, and robust when replicated with a random resampling.

Discussion

Similar to previous research that demonstrated the utility of ANN models for predicting outcomes at the time of hospital discharge, these results demonstrate a fairly high level of accuracy (>85% case classification accuracy; AUCs between .86 and .90) for predicting self-reported ambulation ability 1 year after hospital discharge using demographic and neurologic exam data that are available at the time of admission to SCI rehabilitation. Models for predicting whether individuals with SCI will achieve at least modified independence for self-care activities (vs needing assistance from another person) were moderately accurate (76%−86% case classification accuracy; AUCs between .70 and .82). The performance of models generated by ANNs matched or exceeded the performance of models analyzed using logistic regression constrained by the same independent variables.

Although further studies are needed to prospectively validate the performance of ANN models, the current results support their potential clinical utility for estimating the probability of individual patients achieving at least modified independence for ambulation. In this cohort of individuals with complete and incomplete injuries, activities such as bed-chair transfers, bowel and bladder management, eating, and toileting represent a more complex challenge for predictive modeling than ambulation. This may be because completeness of injury distinguishes prognosis for ambulation ability more clearly than the various predictors of nonambulation outcomes that can be achieved with adaptive equipment in optimal postrehabilitation environments. The predictive models for nonambulation self-care activities depend on a broader range of input variables, including both upper and lower extremity motor scores, and require further optimization to achieve an acceptable level of classification accuracy.

Study limitations

The potential impact of age as a moderator of functional outcomes after SCI is complex,[34] and older age has been found to be a particularly negative prognostic indicator for walking recovery among individuals with AIS C classification.[35] Although a previous predictive model was successfully developed using age 65 as a binary cutoff point,[20] results of other studies[36,37] suggest that dichotomizing subjects at a younger age would be a reasonable consideration. Another limitation of this study is that outcome information about ambulation status relied on subjects' "yes" or "no" responses to 3 questions about walking ability rather than a more finely graded self-report index or an objective measurement

Table 1 Percentage of subjects achieving a modified to complete level of independence for mobility and self-care activities 1 year after hospital discharge as a function of AIS classification at the time of admission to rehabilitation

AIS Classification	Walk 150ft (n=3097)	Walk 1 Block (n=3095)	Walk 1 Flight of Stairs (n=3095)	Self-care Bed-Chair Transfers (n=2919)	Self-care Eating (n=2919)	Self-care Bowel Management (n=2919)	Self-care Bladder Management (n=2919)	Self-care Toileting (n=2919)
A	8.47	6.04	5.33	57.50	74.35	50.29	50.71	47.61
B	23.65	20.69	18.23	51.96	70.76	46.21	46.74	46.48
C	61.67	51.97	52.39	70.58	75.81	68.41	68.41	63.18
D	92.40	85.52	88.04	92.14	91.89	90.62	86.57	89.35

Table 2 Comparison of the accuracy and performance of ANN and logistic regression models for classification of functional outcomes 1 year after hospital discharge

Functional Outcomes	Prevalence (%)	ANN Accuracy (%)	LR Accuracy (%)	ANN PLR	ANN NLR	LR PLR	LR NLR	ANN AUC	ANN SE	ANN Lower 95% CI	ANN Upper 95% CI	LR AUC	LR SE	LR Lower 95% CI	LR Upper 95% CI	*z*	*P*
Walk																	
150ft	42.27	87.74	87.87	8.59	.17	8.43	.16	.8801	.0148	.8510	.9092	.8754	.0067	.8623	.8885	0.289	.0773
1 street block	39.26	85.51	86.02	5.46	.14	6.28	.17	.8874	.0146	.8589	.9159	.8589	.0074	.8444	.8734	1.746	.0808
1 flight stairs	39.10	87.10	87.17	7.57	.17	7.18	.16	.9022	.0137	.8754	.9290	.8688	.0072	.8548	.8828	2.164	.0305
Self-care																	
Bed-chair transfers	68.41	83.37	82.37	2.57	.15	2.60	.15	.8246	.0167	.7919	.8573	.7774	.0084	.7609	.7939	2.525	.0116
Bladder management	63.78	75.93	75.58	1.84	.20	1.84	.20	.7780	.0187	.7414	.8146	.7054	.0095	.6869	.7239	3.467	.0005
Bowel management	64.54	78.31	77.41	2.12	.18	2.07	.20	.8066	.0175	.7724	.8408	.7295	.0091	.7116	.7474	3.912	.0001
Eating	78.50	83.22	85.91	2.12	.17	2.62	.13	.8205	.0179	.7855	.8555	.7828	.0090	.7652	.8004	1.887	.0592
Toileting	61.41	78.47	79.54	2.28	.17	2.35	.15	.8224	.0167	.7896	.8552	.7611	.0086	.7442	.7780	3.261	.0011

Abbreviations: CI, confidence interval; LR, logistic regression.

of walking ability. Additionally, given previous studies[38,39] identifying the relationship between preserved sensory function and subsequent recovery of motor functioning, the absence of sensory data may have detracted from the success of these predictive models. In the interest of allowing a direct comparison with ANN models, the logistic regression analyses in this study were constrained to the same input variables used in the ANN models. Although the performance accuracies of the ANN and regression models were closely parallel, further optimization of unconstrained regression-based algorithms may be possible.

Ethical issues associated with the development and use of predictive analytics in health care have been identified in the ethics codes of various health informatics organizations[40,41] and discussed in specialized forums for medical informatics professionals[42] and policy makers,[43] but have received less attention in the clinical practice literature. The potentially asymmetrical implications of false-positive versus false-negative projections for functional outcomes may be addressed by incorporating misclassification cost analyses into the process of refining predictive models, and the need for continuous improvement is one guiding principle during the development of clinical decision support tools.[17] However, the potential impact of scientifically supported prognostic information on the morale and motivation of individual patients requires health care providers to carefully consider comorbid factors that are not included in the predictive model, to exercise clinical judgment, and to engage patients in the counseling process.

Conclusions

With further validation, such predictive models may allow clinicians to use data available at the time of admission to inpatient SCI rehabilitation to accurately predict longer-term ambulation status, and whether or not individual patients are likely to perform various self-care activities with or without assistance from another person. The use of predictive models to support clinical judgment may reduce the level of uncertainty faced by patients during the early phase of their rehabilitation, provide more accurate individualized estimates about the probability of diverse functional outcomes, and guide recommendations for subsequent rehabilitation.

Suppliers

a. Neuroshell classifier; Ward Systems Group, Inc.
b. IBM-SPSS Modeler; IBM Corp.

Keywords

Activities of daily living; Decision support techniques; Medical informatics; Rehabilitation; Spinal cord injuries

Corresponding author

Timothy Belliveau, PhD, Hospital for Special Care, 2150 Corbin Ave, New Britain, CT 06053. *E-mail address:* tbelliveau@hfsc.org.

References

1. Kirshblum SC, Priebe MM, Ho CH, Scelza WM, Chiodo AE, Wuermser LA. Spinal cord injury medicine. 3. Rehabilitation phase after acute spinal cord injury. Arch Phys Med Rehabil 2007;88(3 Suppl 1):S62-70.
2. Ditunno JF. Predicting recovery after spinal cord injury: a rehabilitation imperative. Arch Phys Med Rehabil 1999;80:361-4.
3. Anderson K. Targeting recovery: priorities of the spinal cord injured population. J Neurotrauma 2004;10:1371-83.
4. Crozier KS, Graziani V, Ditunno JF Jr, Herbison GJ. Spinal cord injury: prognosis for ambulation based on sensory examination in patients who are initially motor complete. Arch Phys Med Rehabil 1991;72:119-21.
5. Crozier KS, Cheng LL, Graziani V, Zorn G, Herbison G, Ditunno JF. Spinal cord injury: prognosis for ambulation based on quadriceps recovery. Spinal Cord 1992;30:762-7.
6. Waters RL, Yakura JS, Adkins RH, Sie I. Recovery following complete paraplegia. Arch Phys Med Rehabil 1992;73:784-9.
7. Waters RL, Adkins RH, Yakura JS, Sie I. Motor and sensory recovery following complete tetraplegia. Arch Phys Med Rehabil 1993;74:242-7.
8. Waters RL, Adkins R, Joy Yakura RP, Vigil D. Prediction of ambulatory performance based on motor scores derived from standards of the American Spinal Injury Association. Arch Phys Med Rehabil 1994;75:756-60.
9. Burns SP, Golding DG, Rolle WA, Graziani V, Ditunno JF. Recovery of ambulation in motor-incomplete tetraplegia. Arch Phys Med Rehabil 1997;78:1169-72.
10. Marino RJ, Ditunno JF, Donovan WH, Maynard F. Neurologic recovery after traumatic spinal cord injury: data from the Model Spinal Cord Injury Systems. Arch Phys Med Rehabil 1999;80:1391-6.
11. Scivoletto G, Tamburella F, Laurenza L, Torre M, Molinari M. Who is going to walk? A review of the factors influencing walking recovery after spinal cord injury. Front Hum Neurosci 2014;8:1-11.
12. Kirshblum SC, Burns SP, Biering-Sorensen F, et al. International standards for neurological classification of spinal cord injury (revised 2011). J Spinal Cord Med 2011;34:535-46.
13. Daverat P, Sibrac MC, Dartigues JF, et al. Early prognostic factors for walking in spinal cord injuries. Spinal Cord 1988;26:255-61.
14. Brown PJ, Marino RJ, Herbison GJ, Ditunno JF Jr. The 72-hour examination as a predictor of recovery in motor complete quadriplegia. Arch Phys Med Rehabil 1991;72:546-8.
15. Consortium for Spinal Cord Medicine. Outcomes following traumatic spinal cord injury: clinical practice guidelines for health-care professionals. Washington (DC): Paralyzed Veterans of America; 1999.
16. Kirshblum SC, O'Connor KC. Levels of spinal cord injury and predictors of neurologic recovery. Phys Med Rehabil Clin N Am 2000;11:1-27.
17. Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE. A roadmap for national action on clinical decision support. J Am Med Inform Assoc 2007;14:141-5.
18. Schopp LH, Hales JW, Brown GD, Quetsch JL. A rationale and training agenda for rehabilitation informatics: roadmap for an emerging discipline. NeuroRehabilitation 2003;18:159-70.
19. Zorner B, Blackenhom W, Dietz V, EM-SCI Study Group, Curt A. Clinical algorithm for improved prediction of ambulation and patient stratification after incomplete spinal cord injury. J Neurotrauma 2010;27:241-52.
20. 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.
21. 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;13:2263-71.
22. 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.

23. Guide for the Uniform Data Set for Medical Rehabilitation (including the FIM™ instrument), version 5.1. Buffalo: State University of New York at Buffalo; 1997.
24. Rowland T, Ohno-Machado L, Ohrn A. Comparison of multiple prediction models for ambulation following spinal cord injury. In: Proceedings of the AMIA Symposium; American Medical Informatics Association; 1998 Nov 7-11; Orlando. Hanley & Belfus, Inc. p 528-32.
25. Gant V, Rodway S, Wyatt J. Artificial neural networks: practical considerations for clinical applications. In: Dybowski R, Gant V, editors. Clinical applications of artificial neural networks. Cambridge: Cambridge Univ Pr; 2001. p 329-56.
26. Abdi H, Valentin D, Edelman B. Neural networks. Thousand Oaks: Sage; 1999.
27. Caudill M, Butler C. Naturally intelligent systems. Cambridge: Massachusetts Institute of Technology; 1990.
28. Muller B, Reinhardt J, Strickland MT. Neural networks: an introduction. 2nd ed. Berlin: Springer-Verlag; 1995.
29. Garson GD. Neural networks: an introductory guide for social scientists. Thousand Oaks: Sage; 1998.
30. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit 1997;30:1145-59.
31. Swets JA. Signal detection theory and ROC analysis in psychology and diagnostics: collected papers. New York: Psychology Pr; 2014.
32. McGee S. Simplifying likelihood ratios. J Gen Intern Med 2002;17:647-50.
33. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin Biochem Rev 2008;29(Suppl 1):S83-7.
34. Furlan JC, Fehlings MG. The impact of age on mortality, impairment, and disability among adults with acute traumatic spinal cord injury. J Neurotrauma 2009;26:1707-17.
35. Scivoletto G, Morganti B, Dittuno P, Ditunno JF, Molinari M. Effects on age on spinal cord lesion patients' rehabilitation. Spinal Cord 2003; 41:457-64.
36. Aito S, D'Andrea M, Werhagen L, et al. Neurological and functional outcome in traumatic central cord syndrome. Spinal Cord 2007;45: 292-7.
37. Cifu DX, Seel R, Kreutzer J, Martwitz J, McKinley W, Wisor D. A multicenter investigation of age-related differences in lengths of stay, hospitalization charges, and outcomes for a matched tetraplegia sample. Arch Phys Med Rehabil 1999;80:733-40.
38. Katoh S, el Masry WS. Motor recovery of patients presenting with motor paralysis and sensory sparing following cervical spinal cord injuries. Paraplegia 1995;33:506-9.
39. Oleson CV, Burns AS, Ditunno JF, Geisler FH, Coleman WP. Prognostic value of pinprick preservation in motor complete, sensory incomplete spinal cord injury. Arch Phys Med Rehabil 2005; 86:988-92.
40. Samuel HW, Zaïane, OR, Sobsey D. Towards a definition of health informatics ethics. In: Proceedings of the 1st ACM International Health Informatics Symposium; 2010 Nov 11-12; Arlington. New York: Association for Computing Machinery; 2010. p 257-64.
41. Goodman KW, Adams S, Berner ES, et al. AMIA's code of professional and ethical conduct. J Am Med Inform Assoc 2013;20: 141-3.
42. Amarasingham R, Audet A, Bates D, et al. Consensus statement on electronic health predictive analytics: a guiding framework to address challenges. EGEMS (Wash DC) 2016;4:1163.
43. Cohen IG, Amarasingham R, Shah A, Xie B, Lo B. The legal and ethical concerns that arise from using complex predictive analytics in health care. Health Aff (Millwood) 2014;33:1139-47.

Supplemental Appendix S1

Methods

Artificial neural networks

ANNs are computing systems that can be used to generate classification or prediction models based on complex nonlinear relationships between variables. The input and output variables of ANNs correspond to the independent and dependent variables in regression analysis, and both methods involve a process by which predictors are optimally weighted. Data analysis using ANNs requires a large archival data set containing known input values (eg, ASIA exam motor scores) and known output classifications (eg, ambulatory status).

These computing systems are composed of an interconnected network of "nodes," which are typically organized into layers. Data enter the network through an input layer. Activation parameters associated with the nodes in 1 or more "hidden" layers are weighted according to mathematical functions. The hidden layers essentially operate as pattern recognition systems, and the resulting classification is forwarded to an output layer. A learning algorithm is optimized as the network is exposed to "training" cases, for which it generates output values (eg, expected scores based on the relationships between input and output values). The network-generated output is compared to the known value of the output variable (ie, datum from the archival dataset). The discrepancy between the network-generated output values and the archive-determined output values is a measurement of the current level of error in the predictive model.[26-28] Many ANN models function through a process of backpropagation (ie, backward propagation of error) in which error data trigger dynamic adjustments to the activation parameters of the weights associated with the nodes.[26,29] The process of generating outputs, revising the error variance, and systematically adjusting the network weights continues through a series of "iterations" or learning "epochs" as the network algorithm successively approximates the known values for the output variable. Network "training" proceeds through a specified subset of cases until a maximal level of classification accuracy is achieved.[26] In order to monitor for possible overfitting of the model to the training dataset, the accuracy and generalizability of the model are established on a cross-validation dataset. Indices of the performance of ANN models include the overall accuracy of correctly classified cases in a contingency table, or use of the area under the receiver operating characteristic curve.[30,31]

Results

Supplemental table S1 shows the performance characteristics of models for predicting subjects' ability to ambulate 150ft at home with or without mobility aids. Inputs consisted of age (dichotomized at 65y), and binary classifications to indicate whether AIS motor scores at L2, L3, and S1 were ≥ 2.

Supplemental table S2 shows the performance characteristics of models for classifying subjects' ability to ambulate 1 street block with or without mobility aids, using the same inputs derived from age, L2, L3, and S1 strength.

Supplemental table S3 shows the performance characteristics of models for classifying subjects' ability to walk up 1 flight of steps with or without mobility aids, using the same inputs derived from age, L2, L3, and S1 strength.

Supplemental table S4 shows the performance characteristics of models for classifying subjects' ability to execute bed-chair transfers at a modified or completely independent level (as measured by FIM ratings of 6 or 7) versus subjects who require any amount of physical assistance from a helper (as measured by FIM ratings <6). Inputs consisted of sex, age (dichotomized at age 65y), lower extremity motor score (categorized in intervals of 10), and 5 binary values to indicate whether motor strength at C5, C6, C7, C8, and T1 was ≥ 2.

Supplemental table S5 shows the performance characteristics of models for classifying subjects' bladder management at a modified or completely independent level (as measured by FIM ratings of 6 or 7) versus subjects who require any amount of physical assistance from a helper (as measured by FIM ratings <6). Inputs consisted of upper extremity motor score and lower extremity motor score (categorized in intervals of 10), and 7 binary values to indicate whether motor strength at C5, C6, C7, C8, L2, L3, and S1 was ≥ 2.

Supplemental table S6 shows the performance characteristics of models for classifying subjects' bowel management at a modified or completely independent level (as measured by FIM ratings of 6 or 7) versus subjects who require any amount of physical assistance from a helper (as measured by FIM ratings <6). Inputs consisted of lower extremity motor score (categorized in intervals of 10) and 5 binary values to indicate whether motor strength at C5, C6, C7, C8, and T1 was ≥ 2.

Supplemental table S7 shows the performance characteristics of models for classifying subjects' ability to eat at a modified or completely independent level (as measured by FIM ratings of 6 or 7) versus subjects who require any amount of physical assistance from a helper (as measured by FIM ratings <6). Inputs consisted of age (dichotomized at age 65y), upper extremity motor score (categorized in intervals of 10), and 4 binary values to indicate whether motor strength at C7, C8, L2, and L3 was ≥ 2.

Supplemental table S8 shows the performance characteristics of models for classifying subjects' toileting ability at a modified or completely independent level (as measured by FIM ratings of 6 or 7) versus subjects who require any amount of physical assistance from a helper (as measured by FIM ratings <6). Inputs consisted of age (dichotomized at age 65y) and 4 binary values to indicate whether motor strength at C7, C8, L2, and L3 was ≥ 2.

Supplemental Table S1 Performance of models for predicting the ability to ambulate 150ft 1 year after hospital discharge

Variables	ANN Training (n=2514)	ANN Cross-validation (n=628)	ANN Sensitivity Analysis (n=3142)	LR (n=3142)
Prevalence	43.32	41.56	42.97	42.97
Classification accuracy	88.03	87.74	87.87	87.87
Sensitivity		84.29	85.19	85.19
Specificity		90.19	89.90	89.90
Positive predictive value		85.94	86.40	86.40
Negative predictive value		88.98	86.96	88.96
AUC		0.8801	0.8883	0.8754

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S2 Performance of models for predicting the ability to ambulate 1 street block 1 year after hospital discharge

Variables	ANN Training (n=2512)	ANN Cross-validation (n=628)	ANN Sensitivity Analyses (n=3140)	LR (n=3140)
Prevalence	37.22	40.60	38.06	38.06
Classification accuracy	86.58	85.51	87.77	86.02
Sensitivity		87.84	88.70	85.38
Specificity		83.91	87.20	86.41
Positive predictive value		78.87	80.98	79.31
Negative predictive value		90.99	92.63	90.64
AUC		0.8874	0.8968	0.8589

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S3 Performance of models for predicting the ability to walk 1 flight of steps 1 year after hospital discharge

Variables	ANN Training (n=2511)	ANN Cross-validation (n=628)	ANN Sensitivity Analysis (n=3140)	LR (n=3140)
Prevalence	37.20	41.56	38.06	38.06
Classification accuracy	88.01	87.10	87.77	87.17
Sensitivity		88.89	88.70	88.07
Specificity		85.83	87.20	85.69
Positive predictive value		81.69	80.98	81.53
Negative predictive value		91.57	92.63	90.92
AUC		0.8998	0.8968	0.8688

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S4 Performance of models for predicting at least a modified independence level of bed-chair transfers, without assistance, 1 year after hospital discharge

Variables	ANN Training (n=2362)	ANN Cross-validation (n=590)	ANN Sensitivity Analysis (n=2952)	LR (n=2952)
Prevalence	68.54	68.30	68.50	68.50
Classification accuracy	83.49	83.37	82.72	82.37
Sensitivity		90.57	90.16	90.25
Specificity		54.71	56.56	65.23
Positive predictive value		84.69	85.43	84.96
Negative predictive value		76.10	75.67	75.47
AUC		0.8246	0.8247	0.7774

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S5 Performance of models for predicting at least modified independence with bladder management, without assistance, 1 year after hospital discharge

Variables	ANN Training (n=2362)	ANN Cross-validation (n=590)	ANN Sensitivity Analysis (n=2952)	LR (n=2952)
Prevalence	62.83	64.41	63.14	63.14
Classification accuracy	76.55	75.93	75.64	75.58
Sensitivity		89.47	90.34	89.70
Specificity		51.43	50.46	51.38
Positive predictive value		76.92	75.75	75.97
Negative predictive value		72.97	75.31	74.43
AUC		0.7780	0.7776	0.7054

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S6 Performance of models for predicting at least a modified independence level of bowel management, without assistance, 1 year after hospital discharge

Variables	ANN Training (n=2362)	ANN Cross-validation (n=590)	ANN Sensitivity Analysis (n=2952)	LR (n=2952)
Prevalence	63.72	65.08	64.00	64.00
Classification accuracy	78.92	78.31	78.32	77.41
Sensitivity		89.32	90.42	88.88
Specificity		57.77	56.82	57.01
Positive predictive value		79.77	78.82	78.60
Negative predictive value		74.38	76.94	74.26
AUC		0.8066	0.8077	0.7295

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S7 Performance of ANN and logistic regression models for predicting at least a modified independence level of self-care while eating, without assistance, 1 year after hospital discharge

Variables	ANN Training (n=2362)	ANN Cross-validation (n=590)	ANN Sensitivity Analysis (n=2952)	LR (n=2952)
Prevalence	79.04	78.14	78.86	78.86
Classification accuracy	87.55	83.22	86.42	85.91
Sensitivity		90.46	92.91	91.49
Specificity		57.36	62.18	65.06
Positive predictive value		88.35	90.16	90.72
Negative predictive value		62.71	70.16	67.22
AUC		0.8205	0.8434	0.7828

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.

Supplemental Table S8 Performance of models for predicting at least a modified independence level of self-care while toileting, without assistance, 1 year after hospital discharge

Variables	ANN Training (n=2362)	ANN Cross-validation (n=590)	ANN Sensitivity Analysis (n=2952)	LR (n=2952)
Prevalence	61.73	61.19	61.62	61.62
Classification accuracy	80.31	78.47	80.01	79.54
Sensitivity		89.75	92.58	90.87
Specificity		60.70	59.84	61.34
Positive predictive value		78.26	78.73	79.05
Negative predictive value		78.98	83.39	80.72
AUC		0.8224	0.8435	0.7611

NOTE. Values are percentages except those for AUC.
Abbreviation: LR, logistic regression.