Corroborating evidence by exploring sources of bias in observational spinal cord injury studies

Marcel A. Kopp, MD,* Peter Martus, PhD,* Ralf Watzlawick, MD, Michael J. DeVivo, DrPH, Yuying Chen, MD, PhD, and Jan M. Schwab, MD, PhD

Correspondence

Dr. Kopp
marcel.kopp@charite.de

Neurology® 2018;91:476–479. doi:10.1212/WNL.0000000000006123

Observational studies investigating large real-life datasets are a valuable resource in clinical research. Understanding the imperfect nature of clinical data, statistical approaches factoring in known confounders are instrumental for rigorously addressing bias.¹ Our recent work identifying pneumonia and postoperative wound infections (Pn/Wi) as risk markers for impaired long-term functional recovery and survival after spinal cord injury (SCI)² was considered as a strong statistical analysis.³ However, some unexplored putative confounders in terms of nonrandom loss to follow-up, temporal changes in clinical practice, and exclusion criteria were discussed.³ In order to evaluate and objectivize for the probability of attrition and temporal and selection bias, we apply and discuss an array of analytical tools extending beyond the format of the original publication.²

Methods

The major analysis framework of the original study regarding FIM_{motor} as primary outcome after SCI constituted a linear mixed regression model after multiple imputation of missing values due to dropout and a sensitivity analysis in complete cases in order to quantify the potential effect of attrition and guide the discussion of the results. This strategy was augmented with models stratified for the American Spinal Cord Injury Association impairment scale (AIS) to circumvent possible interactions with the injury severity.² In the stratified models, the imputed and complete case analyses produced differing results in AIS B and C subgroups.² To address the question whether inaccuracy in the imputations may be a reason for such disparities rather than nonrandom dropout,³ we compare the sociodemographic and neurologic baseline of complete cases with cases lost to FIM_{motor} examinations at 1 or 5 years.

The study included patients enrolled between 1995 and 2005 and the incidence of infections as well as the management of care after SCI may have changed throughout this time frame.³ To reveal temporal differences in the rate of Pn/Wi and evaluate their potential to influence the FIM_{motor} analysis, we compare the Pn/Wi rate by periods of enrollment (1995–1999 vs 2000–2005) and include the enrollment period as an explanatory variable in the linear models.

The exclusion criteria of the study were chosen to prevent unknown comorbidity or concomitant injuries from confounding the analysis. However, exclusion of patients bares the risk of selection bias, which can be explored by comparing characteristics of the included population with those of the “excluded” group. Because the number of excluded patients is relatively small in our study and their functional baseline is frequently missing, an outcome analysis³ would be underpowered. Thus, we analyze the association of sociodemographic baseline characteristics with study exclusion as dependent variable using a logistic regression model. With regard to the exclusion criterion “age <16 or >75 years,” we adjust for age by adding a linear and squared term to the model as we suspected possible effects to be U-shaped with extreme values for small and high age.

*These authors contributed equally to this work.

From Charité-Universitätsmedizin Berlin (M.A.K., R.W., J.M.S.); Berlin Institute of Health, QUEST-Center for Transforming Biomedical Research (M.A.K.); Eberhard Karls Universität Tübingen (P.M.), Germany; University of Alabama at Birmingham (M.J.D., Y.C.); and The Ohio State University (J.M.S.), Wexner Medical Center, Columbus.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Table Comparison of baseline characteristics between the groups “missing (FIM_{motor}) follow-up” vs “complete cases” at 1 year (upper panel) and 5 years after SCI (lower panel)

	Total sample			AIS A			AIS B			AIS C		
	FIM 1 y missing (n = 483)	FIM 1 y complete (n = 720)	p Value	FIM 1 y missing (n = 263)	FIM 1 y complete (n = 368)	p Value	FIM 1 y missing (n = 99)	FIM 1 y Complete (n = 164)	p Value	FIM 1 y missing (n = 121)	FIM 1 y complete (n = 188)	p Value
Age median (IQR)	36 (24–49)	35 (24–47)	0.180	33 (21–46)	30 (22–41)	0.042^a	34 (24–48)	32 (24–45)	0.306	42 (32–53.5)	46 (34.25–56)	0.266
Male n/total n (%)	393/483 (81.4)	571/720 (79.3)	0.380	220/263 (83.7)	287/368 (78.0)	0.078	78/99 (78.8)	131/164 (79.9)	0.832	95/121 (78.5)	153/188 (81.4)	0.536
Caucasian n/total n (%)	281/462 (60.8)	483/698 (69.2)	0.003^a	152/249 (61.0)	258/356 (72.5)	0.003^a	57/96 (59.4)	106/160 (66.3)	0.268	72/117 (61.5)	119/182 (65.4)	0.499
Working n/total n (%)	270/395 (68.4)	533/695 (76.7)	0.003^a	150/215 (69.8)	291/359 (81.1)	0.002^a	52/78 (66.7)	125/158 (79.1)	0.038^a	68/102 (66.7)	117/178 (65.7)	0.873
FIM adm median (IQR)	13 (13–17)	13 (13–17)	0.938	13 (13–14)	13 (13–15)	0.187	13 (13–18)	13 (13–18)	0.402	15 (13–22)	14 (13–21)	0.305
Neurol level median (IQR)	4 (4–5)	5 (4–5)	0.008^a	4 (4–5)	5 (4–5)	0.031^a	5 (4–5)	5 (4–6)	0.454	4 (3–5)	4 (4–5)	0.145
UEMS adm median (IQR)	14 (5–24)	14 (6–25)	0.296	10 (1–22)	12 (4–21.25)	0.132	17 (8.75–25.25)	17 (8.0–29)	0.819	18 (10–29)	17 (9.25–29)	0.605
LEMS adm median (IQR)	0 (0–8)	0 (0–5.25)	0.673	0 (0–0)	0 (0–0)	0.091	0 (0–1.25)	0 (0–0)	0.298	18 (9–32)	20 (8–31.75)	0.568
	Total sample			AIS A			AIS B			AIS C		
	FIM 5 y missing (n = 810)	FIM 5 y complete (n = 393)	p Value	FIM 5 y missing (n = 422)	FIM 5 y complete (n = 209)	p Value	FIM 5 y missing (n = 173)	FIM 5 y complete (n = 90)	p Value	FIM 5 y missing (n = 215)	FIM 5 y complete (n = 94)	p Value
Age median (IQR)	35.5 (24–49)	33 (23–44.5)	0.008^a	33 (22–46)	29 (21–39)	0.001^a	34 (24–47)	31.5 (24–41)	0.261	45 (32–56)	46 (36.75–54)	0.595
Male n/total n (%)	653/810 (80.6)	311/393 (79.1)	0.546	342/422 (81.0)	165/209 (78.9)	0.533	140/173 (80.9)	69/90 (76.7)	0.417	171/215 (79.5)	77/94 (81.9)	0.629
Caucasian n/total n (%)	498/770 (64.7)	266/390 (68.2)	0.231	271/399 (67.9)	139/206 (67.5)	0.912	104/166 (62.7)	59/90 (65.4)	0.644	123/205 (60.0)	68/94 (72.3)	0.039^a
Working n/total n (%)	489/709 (69.0)	314/381 (82.4)	<0.001^a	263/369 (71.3)	178/205 (86.8)	<0.001^a	105/150 (70)	72/86 (83.7)	0.019^a	121/190 (63.7)	64/90 (71.1)	0.220
FIM adm median (IQR)	13 (13–17)	13 (13–17)	0.271	13 (13–14)	13 (13–17)	0.001^a	13 (13–18)	13 (13–18)	0.666	15 (13–22)	13.5 (13–20.25)	0.252
Neurol level median (IQR)	5 (4–5)	5 (4–5)	0.273	4 (4–5)	5 (4–5)	0.068	5 (4–5)	5 (4–6)	0.631	4 (4–5)	4 (4–5)	0.372
UEMS adm median (IQR)	14 (5–24)	14 (6–25.5)	0.134	9 (2–20)	13 (5–23.5)	0.004^a	17 (8.25–27)	16 (8.25–29)	0.679	19.5 (10–30)	15.5 (9–28)	0.114
LEMS adm median (IQR)	0 (0–8)	0 (0–3.75)	0.106	0 (0–0)	0 (0–0)	0.986	0 (0–0)	0 (0–0)	0.514	22 (10–33)	14 (6.5–26)	0.008^a

Abbreviations: AIS = ASIA impairment scale; FIM = functional independence measure motor items; IQR = interquartile range; LEMS = lower extremity motor score; UEMS = upper extremity motor score. The distribution of continuous variables is described as median and quartiles and the Mann-Whitney test is applied for group comparison. Categorical variables are reported as frequencies and percentages and compared using the χ^2 test.

^a P-values <0.05.

The analyses were performed in analogy to the original study² using SPSS for Windows (version 24.0, IBM Corp., Armonk, NY).

Results

Attrition bias

Significant baseline differences between complete cases and cases lost to follow-up exist in the total sample and in all AIS strata, indicating a pattern of attrition not completely at random. Major differences are a lower proportion of Caucasians and patients working prior to injury in some AIS strata with missing FIM_{motor}. Other disparities are concerning age and neurologic or functional baseline (table).

Temporal bias

The Pn/Wi rate significantly decreased over time, from 53.3% (327 of 614 patients) in the period 1995–1999 to 40.2% (237 of 589) in the period 2000–2005 ($p < 0.001$). Adjustment of the linear regression model for enrollment period resulted in a change of only 0.001 in the R^2 without relevant effects on the explanatory value of Pn/Wi regarding variance in FIM_{motor}.

Selection bias

The logistic model reveals that baseline differences between the study population ($n = 1,203$) and excluded patients ($n = 224$) in terms of sex, working status, and education are related to age. The model with linear and quadratic age effects predicts a small exclusion probability of 5%–6% for participants age 33–49 years and an increased risk of 16% for participants age 20 and 62 years (both linear and quadratic term $p < 0.001$). None of the other baseline characteristics remains significant after age adjustment.

Discussion

We corroborated the robustness of the original study design² by reanalyzing the dataset for additional sources of bias. The exploration for nonrandom attrition reveals baseline differences between datasets lost to follow-up and complete cases prevailing in terms of working status and ethnicity explicable by disparities in health care access.⁴ The imputation and the complete case analysis as a sensitivity analysis in the original study differ significantly only in the AIS B and C strata but not in the AIS A stratum or the total sample.² This is most likely connected to clearly smaller sample sizes and higher variance of outcome in AIS B and C groups^{2,5} leading to underpowered strata due to attrition as demonstrated in the explorative analysis. Thus, inaccuracy of the imputation seems unlikely. Regarding temporal bias, the observed decline of infection rates over time is in line with longitudinal trends in the incidence of health care–associated infections.⁶ The adjustment of the regression models for the enrollment period does not reveal temporal bias regarding outcome, similarly to a previous study.⁷ Finally, the risk for selection

bias is addressed by multiple logistic regression models adjusted for the u-shaped distribution of age in the excluded population.

As a common limitation of longitudinal observational trials, unmeasured confounders may exist also in this trial. For example, specific premorbid conditions or the management practice of infections were not documented in the study and the collection of additional data was not possible. Thus, confounding effects related to these items were only addressed indirectly by exclusion criteria or assessment of temporal effects, respectively. In addition, feasible instrumental variables were not available to address unknown confounders beyond the sensitivity analysis. We are fully aware of the fact that multiple imputation of missing values cannot eliminate biases caused by unmeasured confounders. However, we are convinced that imputation should be preferred to a complete case analysis, which additionally contains sources of biases due to measured confounders.

Our work illustrates that the validity of large observational studies, particularly in rare and heterogeneous conditions such as SCI,⁵ can be enhanced by combining comprehensive description of the data with consequent application of statistical tools to explore sources of bias in order to effectively augment the best evidence available.

Author contributions

Marcel A. Kopp: study concept and design, statistical analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for intellectual content. Peter Martus: study concept and design, statistical analysis and interpretation of data, critical revision of the manuscript for intellectual content. Ralf Watzlawick: critical revision of the manuscript for intellectual content. Yuying Chen: study concept and design, critical revision of the manuscript for intellectual content. Michael J. DeVivo: study concept and design, acquisition of data, critical revision of the manuscript for intellectual content. Jan M. Schwab: study concept and design, critical revision of the manuscript for intellectual content.

Study funding

The study received funding from the Wings for Life Spinal Cord Research Foundation, Austria (grant WfL-DE-006/12), the Era-Net-NEURON Program of the European Union (SILENCE #01 EW170A and SCI-Net #01EW1710), and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant 90DP0011 and 90SI0200100). NIDILRR is a Center within the Administration for Community Living, US Department of Health and Human Services.

Disclosure

M. Kopp, P. Martus, R. Watzlawick, M. DeVivo, and Y. Chen report no disclosures relevant to the manuscript. J. Schwab

received funding from the W.E. Hunt and C.M. Miller Endowment and is a Discovery Theme Initiative Scholar (OSU). Go to Neurology.org/N for full disclosures.

Received January 23, 2018. Accepted in final form June 8, 2018.

References

1. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666–676.
2. Kopp MA, Watzlawick R, Martus P, et al. Long-term functional outcome in patients with acquired infections after acute spinal cord injury. *Neurology* 2017;88:892–900.
3. Warner FM, Tong B, Jutzeler CR, Cragg JJ, Scheuren PS, Kramer JLK. Journal Club: long-term functional outcome in patients with acquired infections after acute spinal cord injury. *Neurology* 2017;89:e76–e78.
4. Davis KE. Access to Health Care of Adult Men and Women, Ages 18–64, 2012. Rockville, MD; Statistical Brief (Medical Expenditure Panel Survey [US]); 2001.
5. Fawcett JW, Curt A, Steeves JD, et al. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. *Spinal Cord* 2007;45:190–205.
6. Kanamori H, Weber DJ, DiBiase LM, et al. Longitudinal trends in all healthcare-associated infections through comprehensive hospital-wide surveillance and infection control measures over the past 12 years: substantial burden of healthcare-associated infections outside of intensive care units and “other” types of infection. *Infect Control Hosp Epidemiol* 2015;36:1139–1147.
7. Failli V, Kopp MA, Gericke C, et al. Functional neurological recovery after spinal cord injury is impaired in patients with infections. *Brain* 2012;135:3238–3250.

Are You Ready?

The MACRA’s new Quality Payment Program is now in effect. Learn about your payment and reporting options under the Merit-based Incentive Payment System (MIPS) and Advanced Alternative Payment Models (APMs) and what you need to know to do now to ensure your highest reimbursement in 2019. Learn more at AAN.com/view/QPP and sign up to receive timely updates.

Committed to Making a Difference: 2019 American Academy of Neurology Research Program

The American Academy of Neurology (AAN) is committed to making a profound difference in the lives of researchers, in turn making a difference in the lives of patients with brain disease. The ambitious 2019 AAN Research Program offers opportunities ranging from \$130,000 to \$450,000 and designed for all types of research across all career levels and discovery stages. Pave your own pathway to patient care by applying for one of the opportunities by the October 1, 2018, deadline.

Visit AAN.com/view/ResearchProgram today.

Neurology.org/N Offers Important Information to Patients and Their Families

The *Neurology*® Patient Page provides:

- A critical review of ground-breaking discoveries in neurologic research that are written especially for patients and their families
- Up-to-date patient information about many neurologic diseases
- Links to additional information resources for neurologic patients

All *Neurology* Patient Page articles can be easily downloaded and printed, and may be reproduced to distribute for educational purposes. Click on the ‘Patients’ link on the home page (Neurology.org/N) for a complete index of Patient Pages.