Racial Differences in Data Quality and Completeness: Spinal Cord Injury Model Systems' Experiences

Yuying Chen, MD, PhD,¹ Hui-Yi Lin, PhD,² Tung-Sung Tseng, DrPH, MS, CHES/MCHES,³
Huacong Wen, MD, MS,^{1,4} and Michael J. DeVivo, DrPH¹

¹Department of Physical Medicine and Rehabilitation, University of Alabama at Birmingham, Birmingham, Alabama; ²Biostatistics, School of Public Health, Louisiana State University Health Sciences Center, New Orleans, Louisiana; ³Behavioral & Community Health Sciences, School of Public Health, Louisiana State University Health Sciences Center, New Orleans, Louisiana; ⁴Department of Physical Therapy, School of Health Professionals, University of Alabama at Birmingham, Birmingham, Alabama

Background: Among people with spinal cord injury (SCI), minorities experience a disproportionately higher burden of diseases. Knowledge of data quality by race/ethnicity will help better design racial health disparity research and understand potential errors/biases. **Objective:** To investigate racial/ethnic differences in response completeness in a longitudinal SCI database. **Methods:** This study included 7,507 participants (5,483 non-Hispanic whites, 1,414 non-Hispanic blacks, and 610 Hispanics) enrolled in the National SCI Database who returned for follow-up between 2001 and 2006 and were aged ≥ 18 years at follow-up. Missing data were defined as any missing, unknown, or refusal response to interview items. **Results:** The overall missing rate was 29.7%, 9.5%, 9.7%, 10.7%, 12.0%, and 9.8% for the Craig Handicap Assessment and Reporting Technique-Short Form (CHART) economic self-sufficiency subscale, CAGE questionnaire, drug use, Diener's Satisfaction with Life Scale, Patient Health Questionnaire, and pain severity, respectively. The missing rate for the CHART measure was significantly higher among non-Hispanic blacks and Hispanics than among non-Hispanic whites, after controlling for demographics, injury factors, mode of data collection, and study sites. The missing data in the other outcome measures examined were also significantly higher among non-Hispanic blacks than among non-Hispanic whites but were not significantly different between Hispanics and non-Hispanic whites. **Conclusion:** Our study highlights the importance of research methodology designed to improve non-response or response incompleteness, particularly in non-Hispanic blacks, as we move to reduce racial/ethnic disparities and strive to explain how and why disparities occur in the SCI population. **Key words:** data quality, minority health, spinal cord injuries, surveys and questionnaires

Minorities experience a disproportionately higher incidence, prevalence, mortality, and burden of diseases, and addressing these health disparities has become a national priority in the United States.¹ When examining outcomes of persons with spinal cord injury (SCI), the significance of racial/ethnic (hereafter termed "racial") health disparities is particularly critical. The percentage of racial minorities in the SCI population has increased over the past five decades. Moreover, the incidence of SCI is disproportionately higher among blacks than among whites.² Compared with whites, blacks and Hispanics living with SCI tend to have worse physical and psychosocial outcomes.³⁻⁹ Further

investigation is required to examine the nature of such disparities, explore mechanisms by which these disparities occur, and develop interventions to improve the health of the minority SCI population.

A longitudinal study can be beneficial in health disparity research. Of relevance to the SCI community, a longitudinal study allows the investigation of long-term secondary conditions after SCI by various sociodemographic backgrounds. Mail and telephone survey methodology is particularly appealing in this population, not only for examining subjective well-being of participants, but also for reasons of cost, convenience, and transportation concerns.

Corresponding author: Yuying Chen, MD, PhD, Professor, Department of Physical Medicine and Rehabilitation, University of Alabama at Birmingham, 515 Spain Rehabilitation Center, 1717 6th Avenue South, Birmingham, AL 35249; email: yychen@uab.edu

Top Spinal Cord Inj Rehabil 2018;24(2):110-120
© 2018 Thomas Land Publishers, Inc.
www.thomasland.com

doi: 10.1310/sci2402-110

However, loss to follow-up, non-response, and response incompleteness could skew the results and limit the generalizability to the population of interest.¹⁰⁻¹⁴ Without more research on racial differences in missing data and measurement quality, the current national commitment to reduce health disparities can be compromised.

In the present study, we documented the extent of incompleteness in response to questions of sensitive topics and examined the role of race/ethnicity on data quality among individuals enrolled in a longitudinal SCI database. Multiple factors (personal characteristics, mode of data collection, and study sites) were also taken into account while assessing the racial differences in missing data. This information will inform the design of methods to promote the attainment of high-quality and reliable data from people, regardless of racial background. It will also provide insight into the potential bias in SCI health disparity research that results from racial variation in data quality.

Methods

Study participants

The National Spinal Cord Injury Database (NSCID) captures data from about 6% of new SCI cases in the United States.¹⁵ Since its inception in the early 1970s, 29 federally funded SCI Model Systems (SCIMS) centers have contributed data to the database. Data have been collected on demographics, injury characteristics, medical complications, functional status, and psychosocial outcomes of eligible subjects at initial hospital care

(Form I) and annual evaluations (Form II). Form II follow-up is currently completed at post-injury years 1 and 5 and every 5 years thereafter, until one of the following occurs: death, neurologic recovery, or withdrawal of consent. The inclusion and exclusion criteria, data collection process, quality control, and collaborating centers of the database have been described in detail elsewhere.¹⁵ The Institutional Review Board at each SCIMS center approved data collection and utilization.

The present study included 7,507 NSCID participants from 16 SCIMS centers, who completed the Form II follow-up between March 2001 and September 2006 when various psychosocial outcomes were included in the database. Additional inclusion criteria were (1) age ≥ 18 years at follow-up, (2) discernible degree of neurological deficit, and (3) classification as non-Hispanic white, non-Hispanic black, or Hispanic all races. Non-Hispanic other races (Native American, Asian/Pacific Islander, some other races, and those categorized as multiracial) were not included because they represent only 3% of NSCID participants, which prevents reliable statistics. For those with more than one follow-up during the study period, status at the latest contact was included in the analyses. The final sample size for each post-injury year and racial group is depicted in **Table 1**.

Measures

Data on race/ethnicity were obtained during initial hospital care by asking NSCID participants the following questions, using US Census Bureau guidelines for coding: (1) Do you have a Latino or

Table 1. Sample size by post-injury year

	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30	Total
Sample size, <i>n</i>								
Non-Hispanic white	1,336	1,042	695	588	846	746	230	5,483
Non-Hispanic black	464	325	247	148	129	74	27	1,414
Hispanic	275	136	66	42	43	35	13	610
Total participants	2,075	1,503	1,008	778	1,018	855	270	7,507
Total SCIMS centers	16	16	14	12	11	10	6	16

Note: SCIMS = Spinal Cord Injury Model Systems.

Hispanic background? (2) What race are you? If participants answered yes to the first question, they were assigned to the Hispanic group, regardless of race. Among those without Hispanic background, we distinguished non-Hispanic white and black participants based on their response to the second question.

We examined the missing data rates of the following self-reported measures included in the NSCID: Craig Handicap Assessment and Reporting Technique-Short Form (CHART),¹⁶ CAGE questionnaire,¹⁷ drug use, Diener's Satisfaction with Life Scale (SWLS),¹⁸ Patient Health Questionnaire (PHQ),¹⁹ and pain severity. Missing data were defined as any missing, unknown, or refusal response to the interview items.

CHART was used to measure social participation by six subscales in the database: physical independence, cognitive independence, mobility, occupation, social integration, and economic self-sufficiency. The missing rate for each subscale was 7.9%, 8.4%, 8.5%, 9.3%, 10.8%, and 29.7%, respectively. We selected the economic self-sufficiency subscale, which had the highest missing rate among all subscales, for further analysis. The economic self-sufficiency subscale asks the following multiple-choice questions: (1) Approximately what was the combined annual income, in the last year, of all family members in your household? (2) Approximately how much did you pay last year for medical care expenses?

The CAGE questionnaire was used to identify an alcohol problem by asking the following yes/no questions: (1) Have you ever felt you should Cut down on your drinking? (2) Have people Annoyed you by criticizing your drinking? (3) Have you ever felt bad or Guilty about your drinking? (4) Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover (Eye-opener)? The missing rate for the total CAGE score was 9.5%. Participants were also asked to identify their use of illegal drugs or prescribed medications for non-medical reasons during the past year. The missing rate for drug use was 9.7%.

Subjective overall life satisfaction was assessed with Diener's SWLS, which consists of five statements: (1) in most ways my life is close to my ideal, (2) the conditions of my life are excellent,

(3) I am satisfied with my life, (4) so far I have gotten the important things I want in life, and (5) if I could live my life over, I would change almost nothing. Each is rated on a 7-point Likert-type scale, with responses ranging from *strongly disagree* to *strongly agree*. Because the missing rate was similar across the five items (10.1%, 10.1%, 10.0%, 10.1%, and 10.3%, respectively), the present analysis was limited to the total SWLS score (missing rate, 10.7%).

The 10-item PHQ was used to identify major depressive syndrome and also to indicate the severity of depression, which had missing rates of 10.8% and 12.0%, respectively. We limited the analyses to the measure of severity of depression, which had a higher missing rate. The usual level of pain over the past 4 weeks was also rated by participants, using a 0 to 10 scale. The overall missing rate for pain severity was 9.8%.

Sociodemographic variables, including education, marital status, employment, and primary payer, were updated at each follow-up. Gender and neurologic deficits were assessed during initial hospital care. The mode of data collection was also documented: (1) interview by phone, (2) interview in person, (3) self-administered by mail or in the clinic, and (4) combination of the above methods.

Statistical analysis

The missing rate was calculated for each outcome variable at each post-injury year, stratified by race/ethnicity. The significance of racial differences in missing rates was examined using the chi-square test. Multiple logistic regression was conducted to assess the effect of race/ethnicity on missing data for each outcome variable after accounting for potential confounding factors, including the mode of data collection, age at injury, gender, marital status, education, employment, primary payer, years since injury, and severity of injury. To minimize bias from regional differences in data collection practices, we divided the SCIMS centers into four groups, based on their loss to follow-up rates at post-injury years 1 and 5, and included it as a control variable in multiple logistic regression. The loss to follow-up rates defining these four groups were: (I) year 1, <15%, and year 5, < 30%; (II) year 1, 15%-24%, and year 5, 30%-39%; (III)

year 1, 25%-34%, and year 5, 40%-44%; and (IV) year 1, $\geq 35\%$, and year 5, $\geq 45\%$. Subgroup analyses were also conducted within individual SCIMS centers to examine the effect of regional variation. Because the results of subgroup analyses did not meaningfully change the conclusion about racial differences in data quality, only aggregate data are presented. All statistical analyses were performed using SAS 9.0 (SAS Institute).

Results

Compared with non-Hispanic white (hereafter termed “white”) participants, non-Hispanic black (hereafter termed “black”) and Hispanic participants were more likely to be younger, single, less educated, and unemployed (**Table 2**). The black and Hispanic groups had a higher percentage of Medicaid recipients and a lower percentage of private insurance beneficiaries than did the white group. Although the overall majority of Form II data collection was conducted by phone interviews, face-to-face interviews were more common among blacks and Hispanics than among whites. SCIMS centers with higher loss to follow-up rates seemed to have a higher percentage of Hispanic and black participants than of white participants.

As shown in **Table 3**, CHART economic self-sufficiency questions had the highest missing rate among outcome measures across post-injury years. Missing data seemed to be more problematic during the early years of injury than during the later years (≥ 20 years). Racial differences in missing data were noted for all outcome measures across different post-injury years, although the significance level varied (**Figure 1**). The missing rate for CHART economic self-sufficiency was highest among blacks, followed by Hispanics and finally whites, from post-injury year 1 to year 20 ($p < .05$). A similar trend in missing data (blacks $>$ Hispanics $>$ whites) was observed for all other measures, particularly at post-injury year 5 ($p < .05$). Although it was not statistically significant, Hispanics seemed to have the highest missing rate among the racial groups for all outcome measures at post-injury year 30.

After adjusting for potential confounding factors, the odds of missing data in responses to CHART economic self-sufficiency questions were

about 67% higher among blacks than among whites (odds ratio [OR], 1.67; 95% confidence interval [CI], 1.45-1.92). The odds of missing data for all other outcome measures were similarly and significantly higher among blacks than among whites (**Figure 2**). Individuals of Hispanic origin were about 30% more likely than whites were to have missing data in responses to CHART economic self-sufficiency questions (OR, 1.30; 95% CI, 1.07-1.58), after controlling for confounding factors. The missing rates of other outcome measures did not differ significantly between the white and Hispanic groups ($p > .05$).

Other factors that were significantly associated with missing data for each outcome variable are summarized in **Table 4**. Later post-injury years and a higher education level were associated with a lower missing data rate across all outcome variables. The SCIMS centers that had a higher loss to follow-up rate tended to have a lower missing data rate among respondents for all outcome measures except the CHART economic self-sufficiency scale. For most of the outcome variables, a higher missing rate was more common in males and in people having SCI at older age and more severe injury. Compared with other modes of data collection, the self-administered data collection method associated with a higher missing data rate for the depression variable ($p = .004$), but a lower missing rate for the drug use variable ($p = .056$).

Discussion

The current study examined the significance of racial differences in missing data for various self-reported outcomes among 7,507 individuals with SCI who were registered in a longitudinal SCI database and returned for follow-up assessments between 2001 and 2006. The finding of a higher missing data rate among minorities than among whites is generally consistent with previous studies of recruitment, retention, non-response, or response incompleteness in the general population.^{14,20-23}

Many of the interview questions were personal and private, such as those about household income, alcohol and drug use, and personal feelings. Differential issues of trust by race/ethnicity might make one group less likely than others to respond

Table 2. Sociodemographic and injury characteristics by race/ethnicity

Characteristics [mean $\pm$ SD or <i>n</i> (%)]	Total	Non-Hispanic white	Non-Hispanic black	Hispanic	<i>p</i> ^a
Sample size, <i>n</i>	7,507	5,483	1,414	610	
Current age, years	42.9 $\pm$ 13.6	44.0 $\pm$ 13.5	40.4 $\pm$ 13.7	38.6 $\pm$ 12.7	<.0001
Gender					<.0001
Female	1,521 (20.3)	1,170 (21.5)	231 (16.3)	113 (18.5)	
Male	5,986 (79.7)	4,306 (78.5)	1,183 (83.7)	497 (81.5)	
Education level					<.0001
<High school	1,033 (14.5)	458 (8.8)	373 (27.7)	202 (36.2)	
High school	3,753 (52.7)	2,724 (52.2)	763 (56.7)	266 (47.7)	
>High school	2,338 (32.8)	2,039 (39.1)	209 (15.5)	90 (16.1)	
Unknown	383	262	69	52	
Employment status					<.0001
Working	1,920 (26.1)	1,677 (31.2)	153 (11.0)	90 (15.4)	
Homemaker	109 (1.5)	83 (1.5)	18 (1.3)	8 (1.4)	
Student/Trainee	456 (6.2)	341 (6.3)	71 (5.1)	44 (7.5)	
Retired	697 (9.5)	594 (11.1)	84 (6.1)	19 (3.2)	
Unemployment and others	4,167 (56.7)	2,681 (49.9)	1,061 (76.5)	425 (72.5)	
Unknown	158	107	27	24	
Marital status					<.0001
Single/Never married	2,828 (38.4)	1,831 (34.0)	756 (54.4)	241 (41.1)	
Married	2,646 (36.0)	2,162 (40.2)	275 (19.8)	209 (35.7)	
Divorced, separated, widowed, or others	1,887 (25.6)	1,392 (25.9)	359 (25.8)	136 (23.2)	
Unknown	146	98	24	24	
Primary sponsor of care					<.0001
Private	1,921 (26.1)	1,614 (29.9)	198 (14.4)	109 (18.4)	
Worker compensation	547 (7.4)	426 (7.9)	45 (3.3)	76 (12.8)	
Medicare	2,064 (28.0)	1,564 (29.0)	383 (27.9)	117 (19.7)	
Medicaid	1,579 (21.5)	781 (14.5)	594 (43.3)	204 (34.3)	
Other government insurance	134 (1.8)	97 (1.8)	21 (1.5)	16 (2.7)	
Self-pay, no pay, or other	1,118 (15.2)	916 (17.0)	130 (9.5)	72 (12.1)	
Unknown	144	85	43	16	

(Continued)

Table 2. Sociodemographic and injury characteristics by race/ethnicity (CONT.)

Characteristics [mean $\pm$ SD or n (%)]	Total	Non-Hispanic white	Non-Hispanic black	Hispanic	<i>p</i> ^a
Mode of data collection					<.0001
By phone	5,103 (70.3)	3,768 (70.9)	945 (70.0)	390 (66.4)	
In person	518 (7.1)	282 (5.3)	166 (12.3)	70 (11.9)	
Self-administered	968 (13.3)	764 (14.4)	115 (8.5)	89 (15.2)	
Combination	667 (9.2)	504 (9.5)	125 (9.3)	38 (6.5)	
Unknown	251	165	63	23	
Level and completeness of injury					<.0001
Paraplegia incomplete	1,277 (17.3)	982 (18.2)	185 (13.3)	110 (18.3)	
Paraplegia complete	2,286 (31.0)	1,578 (29.3)	519 (37.4)	189 (31.5)	
Tetraplegia incomplete	2,174 (29.5)	1,625 (30.2)	395 (28.5)	154 (25.7)	
Tetraplegia complete	1,635 (22.2)	1,200 (22.3)	288 (20.8)	147 (24.5)	
Unknown	135	98	27	10	
Model Systems Centers ^b					<.0001
Group I	3,918 (52.2)	2,922 (53.3)	707 (50.0)	289 (47.4)	
Group II	872 (11.6)	779 (14.2)	71 (5.0)	22 (3.6)	
Group III	1,955 (26.0)	1,247 (22.7)	537 (38.0)	171 (28.0)	
Group IV	762 (10.2)	535 (9.8)	99 (7.0)	128 (21.0)	

^a Based on the chi-square test and analysis of variance, as appropriate.

^b Centers grouped by loss to follow-up rates (group I: year 1, < 15%, and year 5, < 30%; group II: year 1, 15%-24%, and year 5, 30%-39%; group III: year 1, 25%-34%, and year 5, 40%-44%; and group IV: year 1, $\geq$ 35% and year 5, $\geq$ 45%).

Table 3. Missing data rate by post-injury year

	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30	Total
Sample size, <i>n</i>	2,075	1,503	1,008	778	1,018	855	270	7,507
Missing rate, %								
CHART	38.3	31.7	28.9	28.0	23.7	18.0	19.3	29.7
CAGE	12.4	9.0	9.7	12.0	7.5	5.4	3.3	9.5
Drug use	12.2	9.0	10.1	12.5	7.6	6.1	3.3	9.7
SWLS	13.4	9.8	10.5	14.7	8.7	6.1	5.2	10.7
PHQ	16.0	11.5	11.5	14.4	9.9	6.7	3.7	12.0
Pain severity	12.6	9.5	9.3	12.1	7.9	6.2	3.3	9.8

Note: CHART = Craig Handicap Assessment and Reporting Technique; CAGE = CAGE questionnaire for alcohol problems; SWLS = Diener's Satisfaction with Life Scale; PHQ = Patient Health Questionnaire.

Figure 1. Missing data rates for self-reported outcome measures in the National Spinal Cord Injury Database by race/ethnicity and post-injury year. Missing data rates were determined for outcome measures captured by Form II survey, including (A) economic self-sufficiency, assessed by the Craig Handicap Assessment and Reporting Technique (CHART), (B) alcohol problems, assessed with the CAGE questionnaire, (C) drug use, (D) satisfaction with life, assessed by Diener's Satisfaction with Life Scale (SWLS), (E) depression, assessed by Patient Health Questionnaire (PHQ), and (F) pain severity. NHW = non-Hispanic white; NHB = non-Hispanic black. * $p < .05$.

to these questions. This trust issue, however, could not be appropriately addressed by the present study because of the constraints of the existing database. Racial matching between data collectors and participants may also contribute

to whether or how participants respond. Previous studies in the general population have shown that the cultural match between interviewers and participants would more easily elicit feelings of trust.^{14,24} Johnson et al also suggested that mistrust

Figure 2. Racial differences in the likelihood of missing data for self-reported outcome measures in the National Spinal Cord Injury Database. Adjusted odds ratios (OR; circles) and 95% confidence intervals (CI, squares) of missing data in outcome measures captured for non-Hispanic black (NHB) participants (left) and Hispanic participants (right) versus non-Hispanic white (NHW) participants. CHART = Craig Handicap Assessment and Reporting Technique; CAGE = CAGE questionnaire for alcohol problems; SWLS = Diener's Satisfaction with Life Scale; PHQ = Patient Health Questionnaire.

in the medical system, lack of racial concordance between physicians and patients, and culturally inappropriate communication would decrease health care participation among minorities.²⁵ Considering that the majority of the SCIMS data collectors are white females, it may be culturally uncomfortable for individuals of racial minorities to answer sensitive questions, particularly during face-to-face interview, which could consequently contribute to refusal. This raises the importance of increasing the diversity and intercultural competency of data collectors to improve their engagement with NSCID participants of different cultures. This finding also supports the necessity of regular monitoring of data quality and providing feedback to data collectors.

Several of these measures have not been validated in diverse groups of the SCI population and may not have cultural relevance to different racial groups, thereby resulting in non-response or incomplete data.^{19,26-28} For example, Dhalla and Kopec reviewed studies of the CAGE questionnaire for alcohol misuse assessment and concluded that

it was an appropriate instrument in medical and surgical inpatients, ambulatory medical patients, and psychiatric inpatients, but did not perform well in primary care patients, white women, prenatal women, and college students.²⁶

This study identified higher missing data rates during the early years of injury. Since post-injury year 1 or even year 5 (if participants missed their first anniversary follow-up) is the first time that study participants are exposed to Form II interview questions, general cooperation may play a key role in determining response completeness or missing data. Individuals who return for follow-up at later years are likely to be more cooperative and willing to answer questions, which contributes to a lower missing data rate regardless of race/ethnicity.

Previous studies suggested that people who have a higher education level might be able to comprehend the survey questions better and as a result are more likely to respond.^{12,13,23} This is supported by the current study of people with SCI, in which we observed an inverse relationship between education level and missing data

Table 4. Personal characteristics associated with missing data: Results from multiple logistic regression^a

Characteristics ^b	CHART	CAGE	Drug	SWLS	PHQ	Pain
Years since injury	0.97*	0.97*	0.97*	0.97*	0.96*	0.97*
Education level						
<High school	Ref	Ref	Ref	Ref	Ref	Ref
High school	0.82*	0.71*	0.68*	0.74*	0.70*	0.66*
>High school	0.48*	0.37*	0.34*	0.40*	0.35*	0.35*
Model Systems Centers						
Group I	Ref	Ref	Ref	Ref	Ref	Ref
Group II	0.72*	0.49*	0.44*	0.54*	0.68*	0.51*
Group III	1.07	0.29*	0.30*	0.29*	0.40*	0.23*
Group IV	1.22*	0.24*	0.24*	0.29*	0.32*	0.40*
Age at injury	1.00	1.01*	1.01	1.01*	1.01*	1.01*
Gender						
Female	Ref	Ref	Ref	Ref	Ref	Ref
Male	1.08	1.41*	1.46*	1.21	1.27*	1.31*
Level and completeness of injury						
Paraplegia incomplete	Ref	Ref	Ref	Ref	Ref	Ref
Paraplegia complete	0.96	1.08	1.14	1.37*	1.07	1.09
Tetraplegia incomplete	1.07	0.91	0.97	1.14	0.90	0.86
Tetraplegia complete	1.38*	1.27	1.29	1.59*	1.48*	1.26
Primary sponsor of care						
Private/worker compensation	Ref	Ref	Ref	Ref	Ref	Ref
Self-pay/other	0.90	1.21	1.09	1.71*	1.20	1.68*
Other government insurance	1.00	1.22	1.22	1.13	1.13	1.20
Mode of data collection						
Not self-administered	Ref	Ref	Ref	Ref	Ref	Ref
Self-administered	1.14	0.78	0.70	0.93	1.45*	1.02
Marital status						
Unmarried	Ref	Ref	Ref	Ref	Ref	Ref
Married	0.70*	0.83	0.83	0.84	0.86	0.81
Employment status						
Others	Ref	Ref	Ref	Ref	Ref	Ref
Work/student/retired/trainee	1.17*	1.02	1.08	1.20	1.16	1.06

Note: Data are presented as odds ratio of missing data, as compared to the reference group. Bold font and asterisk indicate $p < .05$. CHART = Craig Handicap Assessment and Reporting Technique; CAGE = CAGE questionnaire for alcohol problems; SWLS = Diener's Satisfaction with Life Scale; PHQ = Patient Health Questionnaire; Ref = reference.

^aAdjusted for race/ethnicity and listed variables.

^bPersonal characteristics are listed by the order of significance, from being significant for all five outcome measures to only one measure.

rate across all outcome measures. The present finding that SCIMS centers with a higher loss to follow-up rate tended to have lower missing data rates among respondents seems to echo previous studies in the general population that suggested a trade-off between an increased response rate and a decreased response completeness.^{12,23,29} Other center-level factors, such as administrative supports and data collection practices, that are attributable to data quality deserve further research. These control factors seem to well explain the differences in missing data between Hispanics and non-Hispanic whites regarding CAGE, drug use, SWLS, PHQ, and pain measures (**Figure 2**). In contrast, the missing data rate for these five measures remained higher among blacks than among whites after controlling for demographics, injury factors, mode of data collection, and participating SCIMS centers. For CHART economic self-sufficiency questions about household income and medical expenses, a high missing rate (29.7%) is not surprising, given the required knowledge of the facts and sensitivity nature of the questions. However, the differential missing rate by race/ethnicity, higher in both Hispanics and blacks than in whites even after accounting for other factors under investigation, is interesting and deserves further research.

Limitations

The study findings need to be interpreted with caution given the following limitations. The effect of language barrier on missing data and associated bias could not be addressed, given the small number of individuals who were not English or Spanish speakers in the NSCID. Similarly, because of a limited sample size, the other races were excluded from this study. Race/ethnicity reflects many complex and interrelated factors, including socioeconomic status, health behaviors and beliefs, acculturation, racism, and environment.³⁰ The present study addressed some, but not all, of these factors. Finally, readers should note the general constraints of the NSCID that have been described in previous publications, such as hospital-based sampling, loss to follow-up, and other generalizability issues.¹⁵

Conclusion

This is the first study conducted to examine, systematically, racial differences in response incompleteness or missing data in a longitudinal study of people with SCI. As we think about making conclusions regarding racial health disparities, it is important to determine first if there are any significant differences between participants with regard to data quality and how that might affect our conclusions. This study provides the foundation for future investigations that aim to identify best practices and develop training manuals and guidelines on longitudinal data collection among individuals of culturally diverse groups. For example, with the knowledge of a higher missing data rate among blacks, we might ask for an additional review of the interview form for any missing information and further attempt to resolve incomplete data before the close of an interview. The mechanism for handling refusals during data collection, particularly while interviewing minorities, may deserve further discussion and standardization across study sites, as well. In addition, cultural adaptations or racial/ethnic matching of data collectors and prospective participants may encourage participation and improve missing data rates. As we move to reduce racial/ethnic disparities and strive to explain how and why disparities occur in the SCI population, it is critical to improve non-response or response incompleteness, particularly in the minority group. Future studies are also needed to assess whether or how this non-response bias impacts SCI health disparity research.

Acknowledgments

This work is supported by funding from the National Institutes on Disability, Independent Living, and Rehabilitation Research (NIDILRR 90DP0083). 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. The authors declare no conflicts of interest.

REFERENCES

1. US Department of Health and Human Services. *HHS Action Plan to Reduce Racial and Ethnic Health Disparities*. 2011. <https://minorityhealth.hhs.gov/omh/browse.aspx?lvl=2&lvlid=10>. Accessed July 25, 2017.
2. Chen Y, He Y, DeVivo MJ. Changing demographics and injury profile of new traumatic spinal cord injuries in the United States, 1972-2014. *Arch Phys Med Rehabil*. 2016;97(10):1610-1619.
3. Fyffe DC, Deutsch A, Botticello AL, Kirshblum S, Ottenbacher KJ. Racial and ethnic disparities in functioning at discharge and follow-up among patients with motor complete spinal cord injury. *Arch Phys Med Rehabil*. 2014;95(11):2140-2151.
4. Gary KW, Nicholls E, Shamburger A, Stevens LF, Arango-Lasprilla JC. Do racial and ethnic minority patients fare worse after SCI? A critical review of the literature. *NeuroRehabilitation*. 2011;29(3):275-293.
5. Kennedy P, Kilvert A, Hasson L. Ethnicity and rehabilitation outcomes: The Needs Assessment Checklist. *Spinal Cord*. 2015;53(5):334-339.
6. Krause JS, Saladin LK, Adkins RH. Disparities in subjective well-being, participation, and health after spinal cord injury: A 6-year longitudinal study. *NeuroRehabilitation*. 2009;24(1):47-56.
7. Lad SP, Umeano OA, Karikari IO, et al. Racial disparities in outcomes after spinal cord injury. *J Neurotrauma*. 2013;30(6):492-497.
8. Mahmoudi E, Meade MA, Forchheimer MB, Fyffe DC, Krause JS, Tate D. Longitudinal analysis of hospitalization after spinal cord injury: Variation based on race and ethnicity. *Arch Phys Med Rehabil*. 2014;95(11):2158-2166.
9. Meade MA, Lewis A, Jackson MN, Hess DW. Race, employment, and spinal cord injury. *Arch Phys Med Rehabil*. 2004;85(11):1782-1792.
10. Gorey KM, Trevisan M. Secular trends in the United States black/white hypertension prevalence ratio: Potential impact of diminishing response rates. *Am J Epidemiol*. 1998;147(2):95-99; discussion 100-102.
11. Greenland S. Response and follow-up bias in cohort studies. *Am J Epidemiol*. 1977;106(3):184-187.
12. Helasoja V, Prattala R, Dregval L, Pudule I, Kasmel A. Late response and item nonresponse in the Finbalt Health Monitor survey. *Eur J Public Health*. 2002;12(2):117-123.
13. Korkeila K, Suominen S, Ahvenainen J, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol*. 2001;17(11):991-999.
14. Yancey AK, Ortega AN, Kumanyika SK. Effective recruitment and retention of minority research participants. *Annu Rev Public Health*. 2006;27:1-28.
15. Chen Y, DeVivo MJ, Richards JS, SanAgustin TB. Spinal Cord Injury Model Systems: Review of program and national database from 1970 to 2015. *Arch Phys Med Rehabil*. 2016;97(10):1797-1804.
16. Whiteneck GG, Charlifue SW, Gerhart KA, Overholser JD, Richardson GN. Quantifying handicap: A new measure of long-term rehabilitation outcomes. *Arch Phys Med Rehabil*. 1992;73(6):519-526.
17. Ewing JA. Detecting alcoholism. The CAGE questionnaire. *JAMA*. 1984;252(14):1905-1907.
18. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess*. 1985;49(1):71-75.
19. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*. 1999;282(18):1737-1744.
20. Cunningham WE, Hays RD, Burton TM, Kington RS. Health status measurement performance and health status differences by age, ethnicity, and gender: Assessment in the medical outcomes study. *J Health Care Poor Underserved*. 2000;11(1):58-76.
21. Fongwa MN, Cunningham W, Weech-Maldonado R, Gutierrez PR, Hays RD. Comparison of data quality for reports and ratings of ambulatory care by African American and White Medicare managed care enrollees. *J Aging Health*. 2006;18(5):707-721.
22. Nelson KM, Geiger AM, Mangione CM. Racial and ethnic variation in response to mailed and telephone surveys among women in a managed care population. *Ethn Dis*. 2004;14(4):580-583.
23. Steffen AD, Kolonel LN, Nomura AM, Nagamine FS, Monroe KR, Wilkens LR. The effect of multiple mailings on recruitment: The Multiethnic Cohort. *Cancer Epidemiol Biomarkers Prev*. 2008;17(2):447-454.
24. Dotinga A, van den Eijnden RJ, Bosveld W, Garretsen HF. The effect of data collection mode and ethnicity of interviewer on response rates and self-reported alcohol use among Turks and Moroccans in the Netherlands: An experimental study. *Alcohol Alcohol*. 2005;40(3):242-248.
25. Johnson RL, Roter D, Powe NR, Cooper LA. Patient race/ethnicity and quality of patient-physician communication during medical visits. *Am J Publ Health*. 2004;94(12):2084-2090.
26. Dhalla S, Kopec JA. The CAGE questionnaire for alcohol misuse: A review of reliability and validity studies. *Clin Invest Med*. 2007;30(1):33-41.
27. Hall K, Dijkers M, Whiteneck G, Brooks CA, Krause JS. The Craig Handicap Assessment and Reporting Technique (CHART): Metric properties and scoring. *Top Spinal Cord Inj Rehabil*. 1998;4:16-30.
28. Pavot W, Diener E. Review of the Satisfaction With Life Scale. *Psychol Assess*. 1993;5:164-172.
29. Stang A, Jockel KH. Studies with low response proportions may be less biased than studies with high response proportions. *Am J Epidemiol*. 2004;159(2):204-210.
30. Stewart AL, Napoles-Springer AM. Advancing health disparities research: Can we afford to ignore measurement issues? *Med Care*. 2003;41(11):1207-1220.