ELSEVIER

U.S. Military Deployment During 2001–2006: Comparison of Subjective and Objective Data Sources in a Large Prospective Health Study

BESA SMITH, MPH, PHD, DEBORAH L. WINGARD, PHD, MARGARET A.K. RYAN, MD, MPH, CAROLINE A. MACERA, PHD, THOMAS L. PATTERSON, PHD, AND DONALD J. SLYMEN, PHD

PURPOSE: Studies researching service members' health after deployment have relied on self-reported deployment history, although validity of these data remains unknown. This study compared self-reported and electronic deployment data and explored differences in functional health.
METHODS: Self-reported and military deployment data were compared for more than 51,000 participants enrolled in the Millennium Cohort Study (2004–2006). Kappa statistics were used to measure agreement. Analysis of variance was used to assess functional health, as measured by the Medical Outcomes Study Short Form 36-Item Health Survey for Veterans (SF-36V).
RESULTS: Of 51,741 participants who completed the initial deployment question, objective records and self-report agreed in 47,355 (92%). Agreement was substantial for deployment status, frequency, and number of deployments (kappa = 0.81, 0.71, and 0.61, respectively). Deployment start dates agreed within 1 month for 82% of participants confirmed as deployed once. Participants' Mental and Physical Component Summary scores from the SF-36V did not differ by agreement level.
CONCLUSIONS: These findings indicate substantial agreement between self-reported and objective deployment information and no clinically meaningful differences in functional health for the small proportion with inconsistent deployment information. These findings should be reassuring to investigators who examine military deployment as a determinant of future health.
Ann Epidemiol 2007;17:976–982.

KEY WORDS: Health Status, Health Surveys, Military Medicine, Military Personnel, Veterans.

INTRODUCTION

Every major and minor military conflict has historically been accompanied by a research effort to investigate exposures that may have compromised service members' health. The current U.S. military engagements in Afghanistan and Iraq are no exception. Research in the remote past has been based primarily on self-reported deployment data (1–4). Much of the research after the 1991 Gulf War, however, used electronically maintained personnel deployment data created from combat zone pay files and service branch reporting. Anecdotes from some personnel attesting to be deployed but not indicated as such in electronic personnel records, as well as a small number who deny deployment despite objective records, have added to the general belief that 1991 Gulf War electronically maintained deployment data are approximately 95% accurate (4). After the attacks of September 11, 2001, military deployment information was restricted for reasons of security, putting some research efforts on hold and compelling investigators to turn to self-reported deployment histories. Little is known regarding the validity of these data, and no large-scale attempt to measure the accuracy of these data has been reported.

Understanding the strengths and limitations of self-reported deployment history and Department of Defense (DoD) electronically maintained deployment data are important to researchers investigating service members' health after their deployment. The purpose of this study was to conduct a large population-based comparison of self-reported deployment data from the Millennium Cohort Study with DoD electronic deployment data. The latter have recently become available to researchers investigating deployment health-related issues. Additionally, baseline measures of functional health, by level of agreement on deployment history, are described.

METHODS

The Millennium Cohort Study

The Millennium Cohort Study, launched in 2001, is the largest longitudinal study ever undertaken by the DoD.

From the Department of Defense Center for Deployment Health Research, Naval Health Research Center (B.S., M.A.K.R.); Department of Family and Preventive Medicine, University of California, San Diego (B.S., D.L.W., M.A.K.R.); Graduate School of Public Health, San Diego State University (C.A.M., D.J.S.); and Department of Psychiatry, University of California, San Diego (T.L.P.), San Diego, CA.

Address correspondence to: Besa Smith, DoD Center for Deployment Health Research, Naval Health Research Center, P.O. Box 85122, San Diego, CA 92186-5122. Tel.: 619-553-7603; fax: 619-553-7601. E-mail: Besa@nhrc.navy.mil.

Received March 10, 2007; accepted July 21, 2007.

1047-2797/07/$–see front matter
doi:10.1016/j.annepidem.2007.07.102

Selected Abbreviations and Acronyms

DoD = Department of Defense
SF-36V = Medical Outcomes Study Short Form 36-Item Health Survey for Veterans
MCS = Mental Component Summary
PCS = Physical Component Summary
DMDC = Department of Defense Manpower Data Center

The purpose of the study is to evaluate risk factors related to military service that may be associated with long-term health outcomes. A detailed description of the methodology of this study has been reported elsewhere (5, 6). In brief, invited participants were from a stratified random sample of the 2 million U.S. military personnel serving on active duty or in the Reserves or National Guard in October 2000. Women, those with past deployment experience, and members of the Reserve or National Guard were oversampled. There were 77,047 members who completed a baseline questionnaire between 2001 and 2003. The current study used a subset of the 55,021 participants who completed the first follow-up questionnaire between 2004 and 2006 in addition to their baseline questionnaire between 2001 and 2003.

The Medical Outcomes Study Short Form 36-Item Health Survey for Veterans (SF-36V) is a standardized instrument contained within the Millennium Cohort questionnaire. The SF-36V uses standardized scoring algorithms to assess eight scales of health: physical functioning, role limitations caused by physical problems, bodily pain, general health, vitality, social functioning, role limitations caused by emotional problems, and mental health (7–11). If at least half of the questions in a scale were answered, the mean of the score for the complete portion of that scale was used to impute values for the missing questions. Participant responses to questions comprising these eight scales were further condensed into two measures: the Mental Component Summary (MCS) and the Physical Component Summary (PCS) scores. Lower scores correspond to lower levels of health or functioning. The instrument has been found to have high internal consistency in a military population (12). The MCS and PCS scores were used to assess functional health in the present study.

Deployment questions were added to the Millennium Cohort instrument, beginning with questionnaires administered from 2004 through 2006. The initial deployment history question asks "Over the past 3 years, did you receive imminent danger pay, hardship duty pay, or combat zone tax exclusion benefits for deployment to any of the regions listed below?" (17 countries and 5 sea regions, plus fill-in options for regions not listed). Participants also were asked to specify for each location (up to five) the month and year they arrived and departed. To ascertain whether the participant was deployed in excess of five times during the past 3 years, participants are asked "In the past 3 years, have you been to more regions where you received imminent danger pay, hardship duty pay, or combat zone tax exclusion benefits than fit into the space allowed above?"

Participants were defined as deployed or not deployed if they responded "yes" or "no," respectively, to the initial deployment question. Participants defined as having deployed only once completed only one section on the questionnaire for location and deployment dates. Those defined as having deployed more than once completed more than one section for location and deployment dates or responded yes to the question asking if they deployed more than five times over the past 3 years. The study protocol was approved by the Institutional Review Board of the Naval Health Research Center (San Diego, CA).

Defense Manpower Data Center

The Defense Manpower Data Center (DMDC) maintains a database for all deployments. Service members are identified as having deployed by being reported directly from personnel offices of the service branches or based on having received imminent danger pay, hardship duty pay, or combat zone tax exclusion benefits. DoD deployment data include country location code, and start and end dates for each deployment.

Deployment data from DMDC were used to create multiple measures of deployment similar to deployment variables created from Millennium Cohort data. Participants defined as not deployed had no deployments to Iraq or Afghanistan on record up to the submission date of the follow-up questionnaire. Participants defined as having deployed once had no more than one deployment on record before submission of the follow-up questionnaire, and those defined as having multiple deployments had more than one deployment on record with DMDC, with the second deployment beginning before submission of the follow-up questionnaire.

Demographic and occupational data obtained from DMDC included sex, date of birth, education (high school or less, some college/bachelor's degree, advanced degree), marital status (never married, married, no longer married), race and ethnicity combined (white non-Hispanic, black non-Hispanic, other), pay grade (enlisted, officer), service component (active duty, Reserve/Guard), service branch (Army, Air Force, Navy/Coast Guard, Marine Corps), and primary occupational specialty (10 major groups, defined by the *DoD Occupational Conversion Manual*) (13). All demographics reflect status as of follow-up survey submission.

Statistical Analyses

The study population includes all participants who completed the initial yes/no deployment question on the follow-up questionnaire. Agreement between the electronic

AEP Vol. 17, No. 12
December 2007: 976–982

deployment data and the self-reported deployment history on the questionnaire were separated into four categories: participants who self-reported being deployed with confirmation from the electronic data, participants who self-reported being deployed with no electronic data confirmation, participants who self-reported not being deployed but electronic data show at least one deployment, and participants with no evidence of deployment in either subjective or objective data.

The kappa statistic was used to assess the level of agreement between electronic deployment data and self-reported deployment history (14). Defined agreement levels were "greater than substantial" when kappa (κ) was between 0.8 and 1.0, "substantial" (κ = 0.6–0.8), "moderate" (κ = 0.4–0.6), "fair" (κ = 0.2–0.4), and "slight or poor" (κ = 0.0–0.2) (15). A weighted kappa statistic was used to investigate the agreement between data sources for number of deployments. To investigate agreement of deployment dates, only participants identified by both data sources as having deployed were evaluated. Because neither self-report nor electronic deployment data could be considered the gold standard, sensitivity and specificity measures were not calculated.

Agreement between deployment start dates was categorized as follows: within 1 month, within 1–3 months, and greater than 3 months. Only deployment start dates were compared since some participants were still on deployment when they submitted their questionnaire.

Univariate analyses, including *t* tests, were used to assess the significance of unadjusted associations between deployment agreement and functional health. Basic models were fit with and without weights to account for oversampling. An exploratory analysis was conducted to examine regression diagnostics, significant associations, and possible confounding, while simultaneously adjusting for other variables in the model. A manual backward elimination approach was used to investigate confounding. Variables that were not significant at alpha = 0.05 but upon removal distorted the measure of effect by more than 15%, were retained in subsequent modeling. Analysis of variance was used to evaluate the association between functional status and deployment agreement for the MCS, and the PCS. All data analyses were completed using SAS software, version 9.1.3 (SAS Institute, Inc., Cary, NC).

RESULTS

There were 55,021 participants who filled out a follow-up questionnaire between 2004 and 2006. Those who did not complete the initial deployment question (n = 3062), or did not have complete demographic data (n = 214) were excluded from the current study, leaving 51,745 available for the deployment status analysis. Objective records and self-reported deployment were in 92% agreement (n = 47,447) (Table 1). Most agreed they had not deployed before survey submission (n = 32,366, 62.6%), whereas 15,081 agreed they had deployed (29.1%). The remaining 4298 participants (8.3%) self-reported deployment inconsistent with electronic records. There were 3199 participants whose self-reported deployment was not confirmed electronically (6.2%), and 1099 whose electronic deployment could not be confirmed by self-reported data (2.1%). Agreement between electronic and self-reported deployment data was greater than substantial (κ = 0.81).

TABLE 1. Agreement on deployment between self-reported deployment and electronic records for Millennium Cohort participants 2004–2006

	n (% of sample)	kappa
Deployment status (n = 51,745)		0.81
Records agree, not deployed	32,366 (62.6)	
Records agree, deployed	15,081 (29.1)	
Records disagree	4,298 (8.3)	
Deployment frequency (n = 51,582)		0.72*w*[a]
Records agree, not deployed	32,366 (62.8)	
Records agree, deployed once[b]	7,608 (14.8)	
Records agree, deployed more than once[b]	3,109 (6.0)	
Records disagree	8,499 (16.5)	
Deployment start date, deployed once (n = 7,251)		
Records agree within 1 month	5,976 (82.4)	
Records agree within 1 to 3 months	481 (6.6)	
Records discrepant by greater than 3 months	794 (11.0)	

[a]Weighted kappa statistic.
[b]Deployment frequency in the 3 years before survey submission.

Self-reported number of deployments was also compared with electronic records (Table 1). There were 163 participants removed from the analysis who self-reported they deployed but did not complete the section for number of deployments. Of the remaining available for analysis (n = 51,582), most agreed with respect to deployment frequency (n = 43,083, 83.5%), whereas 8499 participants (16.5%) self-reported their deployment frequency differently than was reflected within electronic records. Agreement between the two data sources for deployment frequency was substantial (weighted κ = 0.72). Number of deployments, ranging from 0 to 6 or more, was also compared, for which agreement was moderate (weighted κ = 0.57; data not shown).

Deployment start dates were compared between self-report and electronic data sources. Participants who self-reported a deployment beginning prior to the conception of the electronic deployment database (n = 851), or who self-reported that they deployed but did not provide any dates (n = 446) were removed from the date comparison.

For those individuals whose self-reported and electronic information agreed they had deployed only once (n = 7,251), 82% had deployment start dates agree within 1 month of each other (Table 1). There were 794 individuals whose self-reported deployment start date disagreed with the date in the electronic files by more than 3 months (11.0%). For those whose records agreed they deployed more than once (n = 2698), 50% of participants reported every deployment start date (up to five deployments) within 1 month of dates on file with DMDC (data not shown).

Characteristics of participants were described by level of agreement between self-report and electronic deployment records (Table 2). A greater proportion of those who self-reported deployment that could not be confirmed by electronic records were male, active duty, Army, and combat specialists. Those whose electronic deployment was not confirmed by self-report were proportionately more likely to be enlisted, active duty, Air Force, specialists in electrical and mechanical repair, and have less than a high-school education.

Adjusted models were fit with and without weights to account for oversampling in the SF-36V analyses. However, no discernible differences were found, therefore non-weighted analyses are presented. Participants whose summary score could not be calculated due to insufficient questionnaire responses were removed (n = 834).

MCS scores ranged from 51.8 for confirmed deployers to 52.5 for those who had evidence of deployment in electronic records but did not self-report being deployed (Table 3). On average, the MCS score for confirmed deployers (mean = 51.8) was lower than that for confirmed nondeployers (mean = 52.0) and those who had evidence of deployment in electronic records but did not self-report being deployed (mean = 52.5).

PCS scores ranged from 53.7 for those whose self-report of deployment was not confirmed by deployment records, to 54.8 for those who had evidence of deployment in electronic records but did not self-report being deployed (Table 3). All pairwise comparisons were statistically different from each other. The exception was the mean PCS for confirmed nondeployers (mean = 53.8) was not significantly different from that of self-reported deployers whose data could not be confirmed by electronic records (mean = 53.7). All differences in MCS and PCS scores by agreement group, although statistically different, were extremely small.

DISCUSSION

After the terrorist attacks of September 11, 2001, deployment data maintained by the service branches and the DoD were temporarily restricted for national security reasons. Investigators turned to deployment histories obtained through interview and self-administered questionnaires (1–4), although how well these data sources compare has not been established. Once military electronic records became available again, the current study sought to quantify similarities and differences between self-reported subjective and electronic objective information on deployment.

Data sources were remarkably consistent in identifying deployment. The electronic deployment data source confirmed no deployment for 97% of those who self-reported that they did not deploy, and confirmed deployment for 82% of those who self-reported that they did deploy (κ = 0.81). Agreement was slightly lower, but still substantial (κ = 0.71) for deployment frequency. For those who self-reported that they deployed only once, 70% were confirmed in the electronic data. Among those who indicated that they deployed multiple times, only 42% were confirmed in the electronic data. However, 67% of participants self-reported the majority their deployment dates (up to five deployments) within 1 month of their deployments on record in the electronic deployment data. For these reasons, it is believed that the kappa statistic for deployment frequency could be an underestimation of the true level of agreement.

Incomplete reporting from service branch personnel offices may explain, in part, some of the self-reported deployment data that could not be confirmed electronically. Similarly, participants traveling for training and other deployment-like missions could have misidentified these as deployments even though they do not fit the hazardous duty pay criteria. Finally, it is possible that some individuals misrepresent deployment for secondary gain (16), although this seems less likely to have occurred in a confidential health survey. Regardless of explanation, very few individuals self-reported deployment without objective confirmation.

Interestingly, Air Force members were proportionately more likely to be on file in the military data as having deployed without self-reporting they deployed. It is possible airmen may be less likely to consider themselves as having deployed if their missions, while warranting receipt of hazardous and combat duty pay, originate on U.S. soil or other noncombatant military bases. Individuals may also consider transitions during deployment as continuous deployments, whereas the electronic data reflect these transitions as separate deployments. Some service members who have participated in covert operations or other missions perceived as secret may be unwilling to share deployment information on the questionnaire, though the electronic database may identify them as having deployed. Finally, deployment information passed to DMDC based on platoon level rather than person level could have incorrectly identified whether an individual actually deployed. As with all disagreement between data sources, however, the number found to have objective deployment data without subjective confirmation was very small.

AEP Vol. 17, No. 12
December 2007: 976–982

TABLE 2. Characteristics of Millennium Cohort Study participants (2004–2006) by agreement of deployment data sources

Characteristic[b]	Study sample n = 51,745 %	Deployment agreement[a] Agree not deployed[c] n = 32,366 n (%)	Disagree/ self-report deployment[c] n = 3,199 n (%)	Disagree/ electronic deployment[c] n = 1,099 n (%)	Agree deployed[c] n = 15,081 n (%)
Sex					
Male	73.7	22,225 (68.7)	2,639 (82.5)	847 (77.1)	12,429 (82.4)
Female	26.3	10,141 (31.3)	560 (17.5)	252 (22.9)	2,652 (17.6)
Birth year					
Pre–1960	24.4	9,388 (29.0)	581 (18.2)	227 (20.7)	2,446 (16.2)
1960–1969	40.6	12,928 (39.9)	1,392 (43.5)	454 (41.3)	6,235 (41.3)
1970–1979	30.8	8,914 (27.5)	1,079 (33.7)	363 (33.0)	5,587 (37.1)
1980 forward	4.2	1,136 (3.5)	147 (4.6)	55 (5.0)	813 (5.4)
Education					
High school or less	46.2	13,198 (40.8)	1,511 (47.2)	675 (61.4)	8,506 (56.4)
Some college/bachelor's	39.7	14,028 (43.3)	1,196 (37.4)	357 (32.5)	4,968 (32.9)
Advanced degree	14.1	5,140 (15.9)	492 (15.4)	67 (6.1)	1,607 (10.7)
Marital status					
Never married	17.6	5,429 (16.8)	583 (18.2)	201 (18.3)	2,880 (19.1)
Married	73.3	23,657 (73.1)	2,375 (74.3)	793 (72.2)	11,118 (73.7)
No longer married	9.1	3,280 (10.1)	241 (7.5)	105 (9.5)	1,083 (7.2)
Race/ethnicity					
White non-Hispanic	71.0	23,081 (71.3)	2,199 (68.7)	820 (74.6)	10,632 (70.5)
Black non-Hispanic	12.2	4,199 (13.0)	316 (9.9)	147 (13.4)	1,638 (10.9)
Other	16.8	5,086 (15.7)	684 (21.4)	132 (12.0)	2,811 (18.6)
Military pay grade					
Enlisted	71.6	23,058 (71.2)	2,165 (67.7)	931 (84.7)	10,873 (72.1)
Officer	28.4	9,308 (28.8)	1,034 (32.3)	168 (15.3)	4,208 (27.9)
Service component					
Reserve/National Guard	53.2	18,880 (58.3)	1,244 (38.9)	518 (47.1)	6,865 (45.5)
Active duty	46.8	13,486 (41.7)	1,955 (61.1)	581 (52.9)	8,216 (54.5)
Branch of service					
Army	47.8	15,551 (48.0)	1,634 (51.1)	153 (13.9)	7,400 (49.1)
Air Force	30.2	9,118 (28.2)	759 (23.7)	656 (59.7)	5,073 (33.6)
Navy/Coast Guard	18.1	6,476 (20.0)	633 (19.8)	269 (24.5)	1,960 (13.0)
Marine Corps	4.0	1,221 (3.8)	173 (5.4)	21 (1.9)	648 (4.3)
Occupational category					
Combat specialists	19.2	5,311 (16.4)	877 (27.4)	148 (13.5)	3,590 (23.8)
Electronic repair	8.7	2,668 (8.2)	318 (9.9)	115 (10.5)	1,398 (9.3)
Communications/intel	7.6	2,349 (7.3)	331 (10.4)	73 (6.6)	1,170 (7.8)
Health care specialists	11.2	4,426 (13.7)	220 (6.9)	66 (6.0)	1,067 (7.1)
Other technical	2.4	775 (2.4)	86 (2.7)	24 (2.2)	383 (2.5)
Functional support	22.8	8,415 (26.0)	572 (17.9)	222 (20.2)	2,610 (17.3)
Electrical/mechanic	13.1	3,674 (11.3)	375 (11.7)	307 (27.9)	2,432 (16.1)
Craft workers	2.8	831 (2.6)	59 (1.8)	45 (4.1)	495 (3.3)
Service support	9.7	2,924 (9.0)	276 (8.6)	92 (8.4)	1,721 (11.4)
Trainees, others	2.5	993 (3.1)	85 (2.7)	7 (0.6)	215 (1.4)

[a]All unadjusted associations between deployment agreement and individual characteristics were statistically significant ($p < 0.01$).
[b]Characteristics reflect status as of follow-up survey submission.
[c]Agree not deployed: both self-report and electronic databases reflect no deployment; Disagree/self-report deployment: self-reported deployment but electronic database reflects no deployment; Disagree/electronic deployment: electronic database reflects deployment but participant self-reported no deployment; Agree deployed: both self-report and electronic databases reflect deployment.

Most demographic characteristics were similar between the agreement groups, with the largest differences found between confirmed nondeployers and confirmed deployers. Characteristics proportionately higher in those confirmed deployed by both sources were similar to expected characteristics of recently deployed populations: male, younger, less highly educated, never married, active duty, and combat specialists.

Mental and physical functioning was also investigated to determine whether health characteristics differed by level of agreement. SF-36V MCS and PCS scores are linearly transformed to have a mean of 50 and a standard deviation of 10

TABLE 3. Adjusted means of SF-36V[a] component summary scores among Millennium Cohort participants (2004–2006) by agreement of deployment data sources

		Deployment agreement			
Functional health	Study sample n = 50,748 Mean[c]	Agree not deployed[b] n = 31,879 Mean[c]	Disagree/self-report deployment[b] n = 3,142 Mean[c]	Disagree/electronic deployment[b] n = 1,074 Mean[c]	Agree deployed[b] n = 14,649 Mean[c]
Mental Component Summary	52.5	52.0[1]	51.9[1,2]	52.5[1]	51.8[2]
Physical Component Summary	53.2	53.8[1]	53.7[1]	54.8[2]	54.2[3]

[1, 2, 3] Groups with different numbers have significantly different SF-36V component summary scores using Tukey's adjustment for multiple comparisons ($p < 0.05$).
[a]SF-36V, Medical Outcomes Study Short Form 36-Item Health Survey for Veterans.
[b]Agree not deployed: both self-report and electronic databases reflect no deployment; Disagree/self-report deployment: self-reported deployment but electronic database reflects no deployment; Disagree/electronic deployment: electronic database reflects deployment but participant self-reported no deployment; Agree deployed: both self-report and electronic databases reflect deployment.
[c]Means are adjusted for sex, birth cohort, education, marital status, race/ethnicity, pay grade, service component, service branch, and occupation. Scores are linearly transformed to have a mean of 50 and a standard deviation of 10. Scores higher than 50 reflect better functioning.

to allow comparison between populations (17). As reported elsewhere, Millennium Cohort members exhibited higher unadjusted and weighted mean MCS and PCS scores, implying better health, when compared with the national norms for ages 18–64 years (18) (data not shown). Scores in the current study were similar to the overall Cohort (data not shown) and also above national norms. Though some differences between agreement groups were statistically significant, a difference of five points has been considered clinically and socially meaningful (19), with a 10-point difference considered moderate, and 20 points very large (20). The widest point differential between agreement groups was 1.1 point, implying remarkable consistency in scores across all groups. Interestingly, while differences in mental and physical functioning may be hypothesized to differ between deployers and nondeployers, we found little differences between the groups noted to be deployed and nondeployed by both self-report and military deployment data.

Limitations to these analyses should be noted. The study population consists of a subset of responders to the Millennium Cohort questionnaire and may not be representative of the U.S. military population in general. Multiple metrics have been validated and very little response bias has been identified among Millennium Cohort participants (21, Wells et al., unpublished data) (6, 22–25), but, by design, participants are more likely to be older and include slightly more women than a random sample of current military. The kappa statistic is dependent on the true prevalence of the variable being examined with the statistic tending toward zero as the true prevalence approaches 0 or 1 (26). However, since a considerable percentage of U.S. military service members deployed, this dependence would have an insignificant effect on these findings. Electronic deployment data contain deployments beginning in September 2001. To create an equivalent time comparison between sources, deployments self-reported prior to September 2001 were not considered. Finally, the amount of missing documentation in the electronic deployment data is unknown.

This study has several strengths. To date, the authors are unaware of any other large-scale comparison of self-reported and objective deployment data. Agreement was assessed for deployment status in addition to comparing similarities in reporting deployment start dates. The large sample size allowed for robust comparisons of self-report and objective data, including demographic and health characteristics.

The current study found remarkably strong agreement between self-reported deployment and objective deployment data. Timing and number of deployments also agreed across data sources. There were no clinically meaningful differences in functional health in those whose deployment information disagreed. Although electronic data are currently available to researchers, understanding differences between self-reported and electronic deployment data is imperative in studies of deployment-related health. And occupational exposures of military service, especially in deployment, may be very critical determinants of lifelong health.

We are indebted to the Millennium Cohort Study participants, without whom these analyses would not be possible. We thank Scott L. Seggerman and Greg D. Boyd from the Defense Manpower Data Center, Seaside, California, and Col Karl E. Friedl, PhD, from the U.S. Army Medical Research and Materiel Command, Fort Detrick, Maryland. Additionally, we thank Laura Chu, MPH; Lacy Farnell; Gia Gumbs, MPH; Isabel Jacobson, MPH; Cynthia Leard, MPH; Travis Leleu; Robert Reed, MS; Katherine Snell; Steven Spiegel; Damika Webb; Keri Welch, MA; and James Whitmer, from the Department of Defense Center for Deployment Health Research, and Michelle Stoia from the Naval Health Research Center, San Diego, California. We appreciate the support of the Henry M. Jackson Foundation for the Advancement of Military Medicine, Rockville, Maryland.

Authors' Contributions

B.S. performed the statistical analysis. All authors helped conceive the study, participated in its design and coordination, and helped to draft the manuscript. All authors read and approved the final manuscript.

In addition to the authors, the Millennium Cohort Study Team includes Paul J. Amoroso, MD, MPH (Madigan Army Medical Center, Tacoma, WA); Edward J. Boyko, MD, MPH (Seattle Epidemiologic

Research and Information Center, Veterans Affairs Medical Center, Seattle, WA); Gary D. Gackstetter, PhD, DVM, MPH (Uniformed Services University of the Health Sciences, Bethesda, MD and Analytic Services, Inc. [ANSER], Arlington, VA); Gregory C. Gray, MD, MPH (College of Public Health, University of Iowa, Iowa City, IA); Tomoko I. Hooper, MD, MPH (Uniformed Services University of the Health Sciences, Bethesda, MD); James R. Riddle, DVM, MPH (Air Force Research Laboratory, Wright-Patterson Air Force Base, OH); Tyler C. Smith, MS, PhD (Department of Defense Center for Deployment Health Research at the Naval Health Research Center, San Diego, CA); and Timothy Wells, PhD, DVM, MPH (Air Force Research Laboratory, Wright-Patterson Air Force Base, OH).

This represents report 07-08, supported by the Department of Defense, under Work Unit no. 60002. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of the Army, Department of the Air Force, Department of Defense, Department of Veterans Affairs, the US Government, San Diego State University, or the University of California, San Diego. This research has been conducted in compliance with all applicable federal regulations governing the protection of human subjects in research (Protocol NHRC.2000.007).

REFERENCES

1. Fukuda K, Nisenbaum R, Stewart G, Thompson WW, Robin L, Washko RM, et al. Chronic multisymptom illness affecting Air Force veterans of the Gulf War. JAMA. 1998;280:981–988.
2. Kang HK, Mahan CM, Lee KY, Magee CA, Murphy FM. Illnesses among United States veterans of the Gulf War: a population-based survey of 30,000 veterans. J Occup Environ Med. 2000;42:491–501.
3. Gray GC, Reed RJ, Kaiser KS, Smith TC, Gastanaga VM. The Seabee Health Study: self-reported multi-symptom conditions are common and strongly associated among Gulf War veterans. Am J Epidemiol. 2002;155:1033–1044.
4. Horner RD, Kamins KG, Feussner JR, Grambow SC, Hoff-Lindquist J, Harati Y, et al. Occurrence of amyotrophic lateral sclerosis among Gulf War veterans. Neurology. 2003;61:742–749.
5. Gray GC, Chesbrough KB, Ryan MA, Amoroso P, Boyko EJ, Gackstetter GD, et al. The Millennium Cohort Study: a 21-year prospective cohort study of 140,000 military personnel. Mil Med. 2002;167:483–488.
6. Ryan MA, Smith TC, Smith B, Amoroso P, Boyko EJ, Gray GC, et al. Millennium Cohort: enrollment begins a 21-year contribution to understanding the impact of military service. J Clin Epidemiol. 2007;60:181–191.
7. Ware JE, Kosinski M, Gandek B. SF-36 Health Survey: manual and interpretation guide. Lincoln, RI: Quality Metric Incorporated; 2000.
8. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473–483.
9. Kazis LE, Lee A, Spiro A 3rd, Rogers W, Ren XS, Miller DR, et al. Measurement comparisons of the medical outcomes study and veterans SF-36 health survey. Health Care Financ Rev. 2004;25:43–58.
10. Kazis LE, Miller DR, Clark JA, Skinner KM, Lee A, Ren XS, et al. Improving the response choices on the veterans SF-36 health survey role functioning scales: results from the Veterans Health Study. J Ambul Care Manage. 2004;27:263–280.
11. Kazis LE, Miller DR, Skinner KM, Lee A, Ren XS, Clark JA, et al. Patient-reported measures of health: The Veterans Health Study. J Ambul Care Manage. 2004;27:70–83.
12. Jones D, Kazis L, Lee A, Rogers W, Skinner K, Cassar L, et al. Health status assessments using the Veterans SF-12 and SF-36: methods for evaluating otucomes in the Veterans Health Administration. J Ambul Care Manage. 2001;24:68–86.
13. *DoD Occupational Conversion Manual: Enlisted/Officer/Civilian*. Washington, DC: Department of Defense, Office of the Assistant Secretary of Defense, Force Management and Personnel; 1991.
14. Cohen JA. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.
16. Burkett BG, Whitley G. Stolen Valor: How the Vietnam Generation was Robbed of its Heroes and its History. Dallas: Verity Press, Inc.; 1998.
17. Ware JE Jr. SF-36 Health Survey update. Spine. 2000;25:3130–3139.
18. Ware J, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Lincoln: QualityMetric Incorporated; 2002.
19. Hopman WM, Towheed T, Anastassiades T, Tenenhouse A, Poliquin S, Berger C, et al. Canadian normative data for the SF-36 health survey. Canadian Multicentre Osteoporosis Study Research Group. CMAJ. 2000;163:265–271.
20. Voelker MD, Saag KG, Schwartz DA, Chrischilles E, Clarke WR, Woolson RF, et al. Health-related quality of life in Gulf War era military personnel. Am J Epidemiol. 2002;155:899–907.
21. Smith B, Smith TC, Gray GC, Ryan MAK. When epidemiology meets the Internet: Web-based surveys in the Millennium Cohort Study. Am J Epidemiol 2007; In press.
22. Smith TC, Smith B, Jacobson IG, Corbeil TE, Ryan MAK. Reliability of standard health assessment instruments in a large, population-based cohort study. Ann Epidemiol. In Press.
23. Smith TC, Jacobson IG, Smith B, Hooper TI, Ryan MAK. The occupational role of women in military service: validation of occupation and prevalence of exposures in the Millennium Cohort Study. Int J Environ Health Res. In Press.
24. Smith B, Leard CA, Smith TC, Reed RJ, Ryan MAK. Anthrax vaccination in the Millennium Cohort: validation and measures of health. Am J Prev Med. In Press.
25. Chretien JP, Chu LK, Smith TC, Smith B, Ryan MAK. Demographic and occupational predictors of early response to a mailed invitation to enroll in a longitudinal health study. BMC Med Res Methodol. 2007:7.
26. Thompson WD, Walter SD. A reappraisal of the kappa coefficient. Journal of Clinical Epidemiology. 1988;41:949–958.