

Validation of a Research Case Definition of Gulf War Illness in the 1991 US Military Population

Vincent G. Iannacchione^a Jill A. Dever^a Carla M. Bann^a Kathleen A. Considine^a
Darryl Creel^a Christopher P. Carson^a Heather Best^a Robert W. Haley^b

^aRTI International, Research Triangle Park, N.C., and ^bDivision of Epidemiology, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Tex., USA

Key Words

Epidemiological methods · Case definition · Diagnostic criteria · Confirmatory factor analysis · Health surveys · Persian Gulf syndrome · Validation studies

Abstract

Background: A case definition of Gulf War illness with 3 primary variants, previously developed by factor analysis of symptoms in a US Navy construction battalion and validated in clinic veterans, identified ill veterans with objective abnormalities of brain function. This study tests pretested hypotheses of its external validity. **Methods:** A stratified probability sample (n = 8,020), selected from a sampling frame of the 3.5 million Gulf War era US military veterans, completed a computer-assisted telephone interview survey. Application of the prior factor weights to the subjects' responses generated the case definition. **Results:** The structural equation model of the case definition fit both random halves of the population sample well (root mean-square error of approximation = 0.015). The overall case definition was 3.87 times (95% confidence interval, 2.61–5.74) more prevalent in the deployed than the deployable nondeployed veterans: 3.33 (1.10–10.10) for syndrome variant 1; 5.11 (2.43–10.75) for variant 2, and 4.25 (2.33–7.74) for variant 3. Functional status on

SF-12 was greatly reduced (effect sizes, 1.0–2.0) in veterans meeting the overall and variant case definitions. **Conclusions:** The factor case definition applies to the full Gulf War veteran population and has good characteristics for research.

Copyright © 2011 S. Karger AG, Basel

Introduction

A substantial proportion, perhaps 25% [1], of veterans of the 1991 Persian Gulf War continue to experience a pattern of symptoms that has become known as 'Gulf War illness'. Initial investigations of ill soldiers in units reporting high rates of illness by military medical teams soon after the war documented a list of symptoms, most prominently chronic fatigue, memory/attention problems, personality change and body pain, which began during or soon after the war. Finding little evidence of diagnosable physical or psychiatric illness, including posttraumatic stress disorder, initial medical investigations were unable to define the illness and thus drew no clear associations with environmental conditions in the war [2, 3]. Subsequent medical examinations of tens of thousands of veterans in military and US Department of

Veterans Affairs (VA) Persian Gulf War registries likewise yielded no case definition of the illness and thus no evidence of etiology [4].

Subsequently, a study in a US Navy reserve battalion that served in the Gulf War used principal components factor analysis to identify unique symptom patterns suggesting at least 3 primary syndrome variants comprising a definable Gulf War illness [5]. The syndrome variants were strongly associated with different sets of self-reported environmental exposures [6]; objective testing identified different patterns of altered brain biochemistry and function associated with the syndrome variants [7], and confirmatory factor analysis of the survey questionnaire reproduced the factor structure in a replication sample of primarily US Army veterans [8]. Since the case definition was developed and validated in relatively small samples, however, the acceptance of the case definition has been limited by questions of external validity.

This article describes the design, implementation and primary findings of the US Military Health Survey (USMHS), a computer-assisted telephone interview (CATI) survey designed to test prestated validation hypotheses in a large statistically representative sample of the US military population at the time of the 1991 Gulf War. Evidence that would support its validity would include finding a good fit of the latent factor structure to the symptom data of the Gulf War veteran population, low prevalence of veterans meeting the case definition in the nondeployed military population and substantially higher prevalence in the deployed population, and a strong inverse association with measures of health-related quality of life.

Materials and Methods

The Factor Case Definition

To enable research, one of the authors (R.W.H.) developed a survey questionnaire of typical symptoms of ill Gulf War veterans soon after the war expressly to derive a case definition [5]. Since the illness resembled many psychiatric diseases in being composed of patterns of symptoms without objective signs or laboratory findings, the survey questionnaire was designed to be analyzed by a two-stage principal components factor analysis to resolve ambiguities in common symptom complaints and detect symptom patterns that might represent illness variants linked to specific environmental exposures. Similar approaches have been used to define psychiatric diseases listed in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition [9]. The investigators administered the questionnaire in controlled group settings to 249 members of a naval reserve battalion deployed to the 1991 Gulf War [3]. The analysis yielded evidence of 6 unique symptom patterns suggestive of syndrome variants, and having

any one of these patterns constituted an *overall factor case definition*. With variants 4–6 overlapping variant 2, variants 1–3 were considered *primary syndrome variants* for further study [5].

Syndrome variant 1 ('impaired cognition') was comprised of mild cognitive deficits, including distractibility, forgetfulness, depression and chronic fatigue (daytime sleepiness) – not limiting employment appreciably. Variant 2 ('confusion/ataxia') included reduced intellectual functioning, confusion, vertigo and disorientation, resulting in substantial limitations of employment. Variant 3 ('central neuropathic pain') involved chronic, widespread joint and muscle pains and other sensory abnormalities such as paresthesias and numbness but, as with variant 1, carried little limitation of employment [5].

Epidemiological analysis identified strong associations of the 3 primary syndrome variants with self-reported environmental exposures to different chemical toxins [6]. Repeat administration of the questionnaire to 335 primarily US Army veterans of the 1991 Gulf War replicated the principal component structure, tested by confirmatory factor analysis [8]. Subsequent studies of representative ill veterans and well controls from the naval reserve battalion differentiated the 3 syndrome variants and controls on neuropsychological [10, 11], neurophysiological [10, 12], autonomic [13], brain imaging [7, 14–16] and functional status [17] measures, with abnormalities severest and most widespread in factor syndrome variant 2 (table 1). The 4 subgroups were particularly well differentiated by a discriminant function of changes in regional cerebral blood flow from a cholinergic pharmacological challenge, measured by single-photon emission computed tomography [7] (fig. 1). Of particular note is that the 3 factor syndrome variant groups tended to deviate from the control group in different directions, e.g. syndrome variant 1 being abnormally lower than the controls and syndrome variant 2 higher, so that the composite of the 3 syndrome variant groups would not differ significantly from the controls [7] (fig. 1) – emphasizing the importance of a subclassification of the case definition.

Main Objectives

The USMHS was designed primarily as a confirmatory test of the null hypothesis of no difference in the prevalence rates of the overall factor case definition between the US military personnel deployed to the Kuwaiti Theater of Operations (KTO) during the conflict and those who were medically able but were not deployed to the KTO (the deployable nondeployed). This required estimation of the prevalence of the overall factor case definition, and the individual factor syndrome variant definitions, within a set of predetermined subgroups of interest (reporting domains). The KTO included Saudi Arabia, Iraq, Kuwait, Bahrain, Qatar, United Arab Emirates, and ships in the Persian Gulf. Secondary objectives were to test the fit of a structural equation model of the factor case definition to the survey data, and to assess the association of the case definition with a measure of health-related quality of life.

Sampling Design

The *sampling frame* from which the survey sample was randomly selected was constructed by merging the following two databases:

- (1) The Desert Shield/Storm file and Defense Manpower Data Center (DMDC) Operation Mission/Contingency file (Seaside, Calif., USA) contained one record for each person on ac-

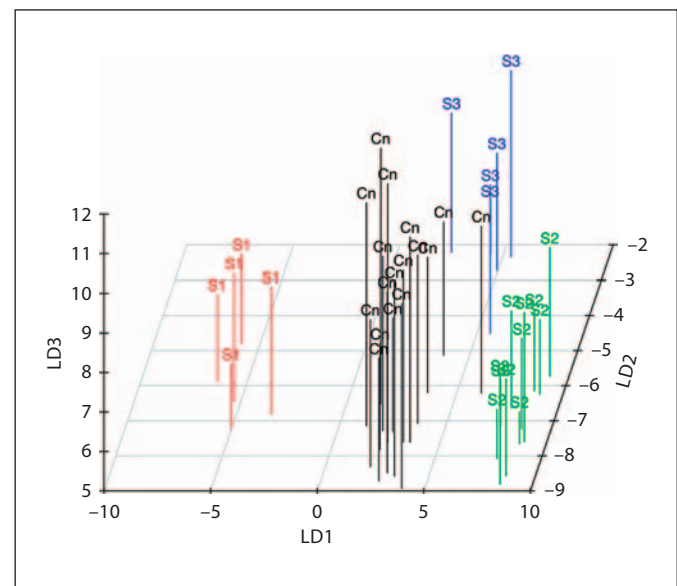
Table 1. Results (means with SEM in parentheses) of medical tests from previously published studies showing differences in overall physical functioning and brain function and metabolism among well veteran controls and the 3 primary syndrome variants defined by the factor case definition

Test	Well veteran controls (n = 20)	Primary factor syndrome variants			p value
		1 (n = 5)	2 (n = 12)	3 (n = 5)	
Health-related quality of life (SF-36) [17]					
Physical component summary	56.1 (1.5)	41.7 (2.6)	31.0 (1.7)	34.4 (2.6)	<0.001
Mental component summary	55.1 (2.4)	46.7 (4.3)	31.7 (2.8)	47.0 (4.3)	<0.001
Circadian variation in parasympathetic nervous system activity (night-day difference in high-frequency heart rate variability) [13]: mean, ms ²	90.9 (21.3)	-6.8 (15.9)	22.9 (15.3)	4.3 (19.0)	<0.001
Chemical analysis of deep brain centers (N-acetylaspartate/creatinine ratio in right basal ganglia) by ¹ H-magnetic resonance spectroscopy) [14]: mean ratio	4.08 (0.13)	3.95 (0.24)	3.35 (0.11)	3.90 (0.18)	0.003
Integrity of acetylcholine receptors in brain (response of regional cerebral blood flow to physostigmine challenge measured by SPECT brain scan, least significant interval) [7]: mean difference in rCBF, ml/mg/min	-1.43 (2.93)	-4.30 (5.25)	4.26 (3.5)	-4.56 (5.17)	0.005

Results are data from a nested case-control study of 22 cases and 20 age-sex-education-matched controls selected from a 1995 epidemiological survey of 249 members of a Naval Reserve construction battalion deployed in the combat zone of the 1991 Persian Gulf War [5]. The studies were performed in 1998 with the subjects residing in the General Clinical Research Center at

the University of Texas Southwestern Medical Center, Dallas, Tex., USA. p values from 4-group ANOVA. The physical and mental component summary scores were calculated from the 8 SF-36 scales as previously published [53]. Values are T scores with the mean of the 1998 US population approximately 50 and SD 10.

Fig. 1. Results of a previously published [7] linear discriminant analysis to identify a subset of the deep brain regions whose mean normalized regional cerebral blood flow measured by single-photon emission computed tomography scanning under the baseline or physostigmine-stimulated condition would jointly classify subjects into the 4 clinical groups defined by the factor case definition. The discriminant model of normalized regional cerebral blood flow from 17 brain regions from either the baseline session or the physostigmine-stimulated session yielded 3 linear discriminant functions (LD1–LD3) that best classified the subjects. The 3 linear discriminant functions provided clear separation of the 4 groups, with factor syndrome variant 1 in red, factor syndrome variant 2 in green, factor syndrome variant 3 in blue, and the control (Cn) group in black. The 3 primary syndrome variants had brain imaging abnormalities deviating from the control group in different directions so that the composite of the 3 syndrome variant groups would not differ significantly from the controls, emphasizing the importance of a subclassification of the case definition. Reproduced with permission from *Psychiatry Research and Neuroimaging* [7].



Downloaded from <http://karger.com/heap/article-pdf/37/2/129/3153937/000331478.pdf> by guest on 07 December 2023

- tive duty, in the reserves, and in the National Guard on August 2, 1990 [18, 19]. This file included historical data and updated information on characteristics such as decedent status, current military status, and last known residence or duty station.
- (2) The US Army Services Center for Unit Records Research database contained records of the geographic location (longitude and latitude) of most military units that served in the Gulf War for each day during the conflict period and afterward.

Under a research protocol approved by the institutional review boards of the US Army and the authors' research institutions, DMDC provided the personally identifying contact information for only the members of the final survey sample. A certificate of confidentiality protecting the privacy of the survey participants was obtained from the National Institute of Environmental Health Sciences prior to the start of data collection.

The *inferential population* for the USMHS comprised all Gulf War era veterans who were living in the 50 United States and Washington, D.C., at the time of data collection and who were physically and mentally able to complete a telephone interview. The inferential population was divided into two subpopulations:

- The *deployed* subpopulation consisted of all active-duty and ready-reserve military personnel (including Coast Guard) who served in the KTO any time from August 2, 1990, through July 31, 1991. This was defined by the binary deployment flag in DMDC's Desert Shield/Storm file updated by a series of deployment questions in the CATI.
- The *deployable nondeployed* subpopulation consisted of the complement of active-duty and ready-reserve personnel serving in August 1990 in any location other than the KTO, excluding persons who were not deployable because of illness (*medically nondeployable*). Medically nondeployable personnel – identified during the interview from questions on illnesses, other than pregnancy, in the 2 years before the war that precluded deployment – were excluded from the referent population to avoid the 'healthy-warrior effect' [20–23].

Allocation and Selection of the Sample

The sample was allocated to ensure adequate numbers of observations in the domains hypothesized to be associated with symptoms of Gulf War illness. The frame was stratified into 229 sampling strata by crossing the variables in each of the following three major strata.

(1) Not deployed to KTO:

- Age group as of January 1, 2007 (<49, ≥49 years) [5, 18, 24]
- Gender [18, 24]
- Race/ethnicity (non-Hispanic White, other race/ethnicity) [18, 25]
- Military component (active duty, reserve/National Guard) [18, 24, 26–28]
- Military occupation (air flight crew, aircraft maintenance, army special operations, other) [29]

(2) Deployed to KTO:

- Location on January 20, 1991 [18, 24]
- Age group as of January 1, 2007 (<49, ≥49 years) [5, 18, 24]
- Gender [18, 24]
- Race/ethnicity (non-Hispanic White, Black/other) [18, 25]
- Military component (active duty, reserve/National Guard) [18, 24, 26–28]

- Military occupation (air flight crew, aircraft maintenance, army special operations, other) [29]
- Stationed at Camp Doha, Kuwait, between July and November 1991 [30]

(3) Groups for special studies:

- Member of a twin pair (one or both siblings deployed to KTO)
- Member of 24th Reserve Naval Mobile Construction Battalion (Seabees) [5, 24]
- Parent of a child with Goldenhar complex birth defect [31]

To test the confirmatory hypothesis, sample size requirements (before and after attrition) were estimated to detect a difference in syndrome prevalence of 5 percentage points for the domains within the deployed population and 10–15 percentage points for comparisons between deployed and deployable nondeployed domains at a one-tailed significance level of 5% with 80% power. Application of a one-tailed test was justified by the prestated confirmatory purpose of the survey and the findings of all prior surveys of Gulf War era veterans, including the pilot phase of this survey, of higher symptom rates in deployed than nondeployed samples [18, 26–28]. Estimated prevalence rates for the domains used in the sample allocation were obtained from prior studies [32] and based on illnesses with definitions closely associated with components of the factor syndromes, e.g. symptoms of fibromyalgia with prevalence estimates of 18–24% in the deployed populations versus 9–13% in nondeployed veterans.

To allow for the increased efficiency of hypothesis testing with the planned multivariable analysis, a logistic regression analysis of pilot survey data was performed to predict the overall factor case definition adjusting for age, gender, race/ethnicity and active/reserve status resulting in an R^2 of 0.12. Using this result, a compression factor of 0.88 ($1 - R^2$) was applied to the expected variances of the prevalence rates to reflect the expected gain in precision produced by the model [33].

The final allocation of the sample among the strata was optimized with the Sample Planning Tool software developed by RTI for DMDC [34]. This software uses a nonlinear algorithm satisfying the Karush-Kuhn-Tucker necessary conditions [35] for optimally minimizing the variable costs of data collection subject to constraints set on the precision of the key survey estimates. The data collection cost model was expressed as a linear, convex function while the equality constraints were defined with respect to the sample design through a set of concave functions. These parameters provide the sufficient conditions needed to ensure an optimal allocation for the USMHS [36, 37]. In addition to the survey design, the precision was conditioned on the actual stratification affected by unequal stratum weighting (which increases variances of final parameter estimates) and sample inefficiencies associated with nonlocation and nonresponse. After inflating the allocation solution for expected response rates, a stratified random sample of 14,817 Gulf War era veterans was selected (fig. 2). Sample members were selected with equal probabilities within each design stratum. All members of the Seabees battalion and parents of Goldenhar children were selected for the study (certainty strata).

Questionnaire Content

The CATI questionnaire comprised three modules administered to all participants in the following order:

- (1) The *Symptoms* module included the questions for constructing the syndrome variant factors and overall factor case definition [5, 8], as well as supplementary symptom information for comparison with other research case definitions (i.e. CDC multisymptom illness [38] and modified Kansas [18] definitions) and similar conditions (e.g. chronic fatigue syndrome, fibromyalgia). Medically nondeployable personnel were identified during the interview from questions on prewar illnesses that had precluded their deployment and were excluded from the referent population to avoid a 'healthy-warrior effect' [20–23].
- (2) The *Exposure* module measured environmental and other risk factors related to Gulf War illness. The locations of persons deployed to the KTO were determined to ensure that the effects of exposure to areas with suspected chemical warfare releases could be evaluated. Questionnaire skip patterns were used to avoid asking nondeployed participants questions about exposures encountered only during deployment.
- (3) The *Family Issues* module covered health issues of the respondents' families that could possibly have resulted from war-related chemical exposures. For example, questions were included to ascertain the numbers of pregnancies, miscarriages, stillbirths, live births, birth defects and learning disabilities in offspring conceived by or born to the subjects and their partners, as well as problems with infertility.

Data Collection

Interviewing for a pilot survey of 200 veterans randomly selected from the target population occurred in 2005–2006 to test the CATI content and interviewing procedures and to generate parameters used to estimate the main study sample size. Telephone interviewing for the full USMHS ran from May 22, 2007, through April 26, 2009, with a dormant period of no outbound calling between June 1 and October 27, 2008.

Current telephone numbers of sampled veterans were initially sought by batch tracing of name, birth date and social security number through the National Change of Address file and other online resources. Unlocated sample members were periodically traced through an interactive search of multiple open and commercial sources. The final unlocatable veterans were sought with addresses submitted with 2007 US income tax returns provided by the Internal Revenue Service.

Sample members with a locatable address were mailed a packet that included the purpose of the study, the importance of their responses, the voluntary nature of their participation, materials to facilitate the interview, an endorsement letter from the American Legion, the internet address of a project website containing additional background information, a 10-dollar bill, and a promised USD 40 upon completion of the interview. Because federal funding of the survey precluded offering financial incentives to the 6% of sample members still on active military duty, they were offered a study-engraved pen and keychain.

Interviewers, certified after a 4-day training course, contacted and interviewed sample members by telephone, primarily during evening and weekend hours. Average interview times varied from approximately 60 min for the nondeployed nonill to 2.5 h for the deployed ill. All participants were informed about the usual length of the interview and were offered multiple sessions if they became fatigued. Since the symptom questions appeared first in the interview, they should not have been affected by interview

Table 2. Effective sample sizes required for testing the association of the overall factor case definition with deployment

Reporting domain	Deployable nondeployed			Deployed		
	target	actual	difference	target	actual	difference
Males ¹	111	591	480	111	2,903	2,792
Females ¹	200	174	–26	200	539	339
Age <49 ¹	111	469	358	111	1,123	1,012
Age ≥49 ¹	200	383	183	200	534	334
Non-Hispanic white ¹	111	537	426	111	819	708
Black/other ¹	298	304	6	298	344	46
Active duty ¹	111	422	311	111	2,597	2,486
Reservists ¹	200	374	174	200	1,027	827
Flight crews ²	46	56	10	46	153	107
Aircraft maintenance ²	46	58	12	46	198	152
Army special forces ²	46	49	3	46	86	40

Effective sample sizes: required for comparisons of the overall factor case definition between deployed and deployable nondeployed reporting domains at the 0.05 one-tailed significance level with 80% power; the target effective sample sizes were based on the USMHS Pilot Survey.

¹ Detectable difference of 10%. ² Detectable difference of 15%.

length or continuation sessions. Only after initial refusal conversion techniques had failed, such as stressing the importance of the study, were non-active-duty sample members offered an additional USD 25 for a total of USD 65 to complete the survey.

Statistical Power

A total of 8,020 persons, who completed at least the *Symptoms* module, were included in the analysis file, constituting an overall response rate of 60.1%, using the AAPOR response rate RR4 definition [39] (fig. 2). The effective number of respondents required to test the main hypothesis, as estimated during the design phase of the study, was exceeded in all domains except for deployed females (table 2). (The effective sample size is the estimated sample size adjusted for unequal weighting. It can be thought of as the sample size equivalent to that drawn by simple random sampling.) In spite of the shortfall, the desired power to detect a difference in the overall factor case definition between the deployed and nondeployed females was achieved because of the larger than expected number of nondeployed female respondents.

Construction of Analysis Weights for Bias Reduction

We developed an analysis weight variable to correct bias from nonproportional sampling of strata and from inability to locate (nonlocation) or obtain participation (noncooperation) from veterans selected into the sample. The ability to locate and to obtain consent for study participation from subjects exploits different processes in most household interview surveys [40]. The USMHS process for locating veterans was largely dependent on the success of the tracing activities. By contrast, the likelihood of participa-

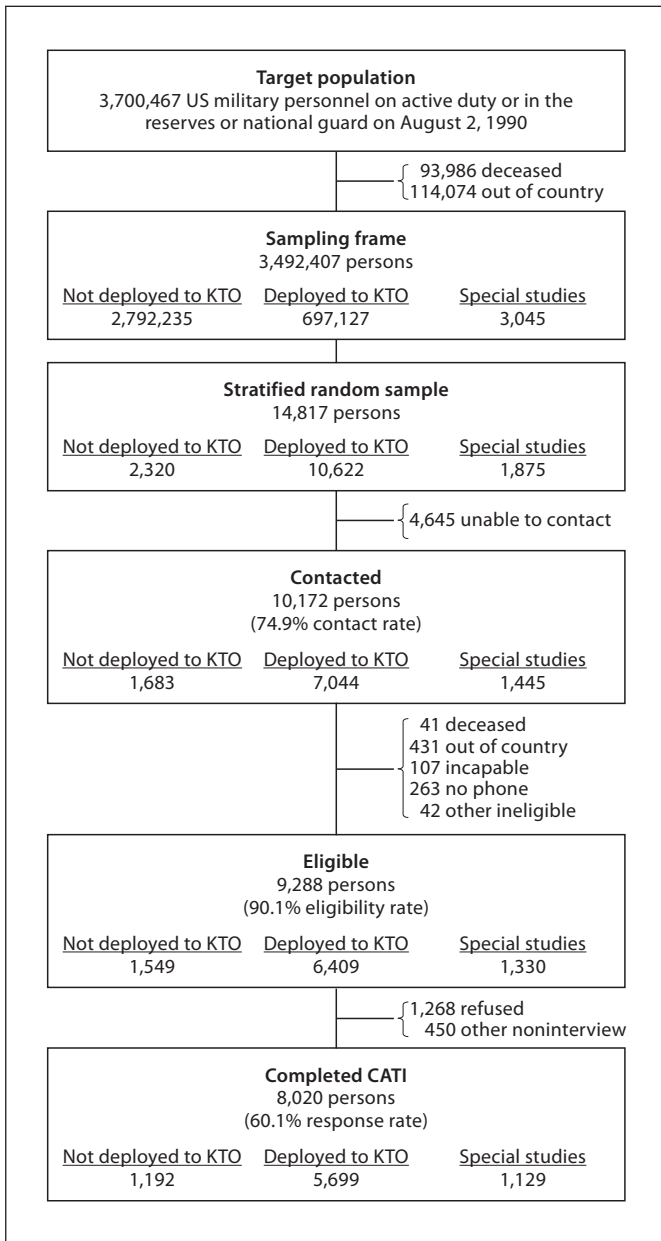


Fig. 2. Sample selection process for the USMHS. ‘Not deployed to KTO’ includes medically nondeployable personnel. ‘Special studies’ included twin pairs, members of the 24th Reserve Naval Construction Battalion (Seabees) and parents of children with Goldenhar complex. Counts for subgroups are suppressed to maintain confidentiality according to terms of the Certificate of Confidentiality. The ‘contact rate’ includes in the base the number of known eligible cases and the estimated number of eligible cases among the undetermined cases. The ‘eligibility rate’ is among sample members with known survey eligibility. The ‘response rate’ is the American Association for Public Opinion Research Response Rate 4 (RR4) and includes in the base the estimated number of eligible cases among those initially selected for the CATI phase of the study [39].

tion was likely affected by numerous factors including the sampled veterans’ military experience. As a result, we constructed survey analysis weights by combining adjustment factors for the sampling design, nonlocation and nonparticipation by the following five-step process [41].

Step 1. A survey design weight was calculated for each sample member as the inverse of the selection probability within the respective design stratum. The design weights had the following form:

$$d_{hi} = \frac{N_h}{n_h} \quad (1)$$

where n_h is the number of sample members selected within stratum h ($h = 1, \dots, 229$) and N_h is the total number of veterans on the sampling frame within stratum h . The design weight d_{hi} is the same for every sample member in the same sampling stratum.

Step 2. The design weights were first adjusted to minimize bias associated with nonlocation using adjustment classes defined by a classification and regression trees algorithm [42] and applied to data available on both located and nonlocated sample veterans. The classes were formed using variables such as service, service component and age group in addition to certain paradata variables (the complete list of variables has been withheld in accordance with conditions in the Certificate of Confidentiality). The adjustments ($\hat{\lambda}_{hi}$) can be written in terms of a logistic model containing the classification and regression tree adjustment classes:

$$\hat{\lambda}_{hi} = P[L_{hi} = 1 | d_{hi}, \mathbf{X}_{hi}, \hat{\beta}_1] = \left[1 + \exp(d_{hi} \mathbf{X}'_{hi} \hat{\beta}_1) \right]^{-1}$$

where d_{hi} is the design weight given in equation 1, \mathbf{X}_{hi} ($l \times 1$) is a vector of indicator variables that identify membership in one of the l adjustment classes, $\hat{\beta}_1$ ($l \times 1$) is a vector of estimated model parameters, and $L_{hi} = 1$ if sample member hi was located (zero otherwise). The resulting location-adjusted analysis weight is written as:

$$w_{1hi} = \begin{cases} d_{hi} \hat{\lambda}_{hi} & \text{for located sample members} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Step 3. A subsequent adjustment was applied to the weights to address any potential nonparticipation bias among sample members who were contacted. Procedures similar to those discussed in step 2 resulted in the following weight adjustment:

$$\hat{\rho}_{hi} = P[P_{hi} = 1 | w_{1hi}, \mathbf{Z}_{hi}, \hat{\beta}_2] = \left[1 + \exp(w_{1hi} \mathbf{Z}'_{hi} \hat{\beta}_2) \right]^{-1}$$

where w_{1hi} is the adjusted weight given in equation 2, \mathbf{Z}_{hi} ($p \times 1$) is a vector of indicator variables that identify membership in one of the p adjustment classes, $\hat{\beta}_2$ ($p \times 1$) is a vector of estimated model parameters, and $P_{hi} = 1$ if sample member hi was located and participated in the USMHS (zero otherwise). The weight adjusted for nonlocation and nonparticipation was then computed as:

$$w_{2hi} = \begin{cases} w_{1hi} \hat{\rho}_{hi} & \text{for located sample members who} \\ & \text{participate in the USMHS} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Step 4. Extreme values of w_{2hi} (falling outside the median $\pm 2 \times$ interquartile range) were trimmed to ensure that excessive variation in the weights would not unnecessarily degrade the precision of the survey estimates.

Step 5. The trimmed weights were ratio adjusted to sum to the number of persons on the sampling frame within the key reporting domains (table 2). The final USMHS analysis weight was constructed as:

$$w_{3hi} = w_{2hi} a_{1hi} a_{2hi} \quad (4)$$

where a_{1hi} is the weight trimming adjustment discussed in step 4 ($a_{1hi} = 1$ for most responding sample members), and $a_{2hi} = f(w_{2hi} a_{1hi})$, the poststratification adjustment calculated as a function of the trimmed weights.

Classification of Syndromic versus Nonsyndromic

The 6 syndrome factor scales were generated by summing the responses to symptom questions multiplied by the scoring weights from the original exploratory factor analysis [5, 8]. The resulting factor scales were then dichotomized, as before, at 1.5 standard deviations where values 1.5 and above were classified as *syndromic* [5, 8]. The cutpoint of 1.5 standard deviations, originally selected from inspection of the scale distributions only [5, 8], was later found to identify syndromic groups with severer functional disability on the Medical Outcomes Study 36-Item Short Form (MOS SF-36) physical component score and mental component score [17]. Persons met the overall factor case definition if they met any of the 6 dichotomized component definitions.

Confirmatory Factor Analysis

A cross-validation approach was used to split the sample (excluding the special samples of twins, Goldenhar parents and Seabees) into two halves. The halves were selected using stratified random sampling to ensure adequate representation of the sampling strata in each half. Confirmatory factor analysis, performed with the M-Plus software [43] incorporated the sampling strata and analysis weights.

Statistical Analyses

Individual subjects' scores on the MOS SF-12 physical and mental summary scales (version 1) were calculated from the 12 questionnaire items by the standard Medical Outcome Trust's scoring algorithm, using the May 2006 SAS program (Janel Hammer), checked against scores published in the 2001 nationally representative sample of the noninstitutionalized general US population in the Medical Expenditure Panel Survey (<http://meps.ahrq.gov>) [44]. Statistical analyses of the survey data were performed with SUDAAN® programs [45] allowing for the complex stratified random sampling design and applying the analysis weights to adjust for unequal selection probabilities and minimize non-response bias. Odds ratios and standard errors of the association of the case definition with deployment were obtained with SUDAAN proc rlogist, and mean SF-12 summary scores in clinical groups defined by the case definition were obtained with SUDAAN proc regress – both analyses controlling for age, gender and race/ethnicity. Seventy-one nondeployed respondents, found by their questionnaire responses to have been medically nondeployable, were excluded from all analyses to avoid bias from the 'healthy-warrior effect' [20–23].

Results

External Validity of the Factor Case Definition

The confirmatory factor analysis found that a parsimonious structural equation model, previously developed to express the 3 primary syndrome factors and a second-order overall Gulf War illness [8] (fig. 3), fit the two random split halves of the survey database ($n = 3,408$ each) well and showed invariance of fit (forced equal loadings) across the two halves (table 3). The goodness-of-fit statistics indicated that the model fit the two random halves of the survey database as well as the same model fit a previous validation sample of US Army veterans recruited from the clinic of a VA medical center [8] (table 3).

Association with Deployment

The prevalence of illness by the overall factor case definition was 3.98% in the nondeployed and 13.59% in the deployed, for a deployment odds ratio adjusted for age, gender and race/ethnicity of 3.87 (95% confidence interval, 2.61–5.74). The deployment odds ratio was highest for factor syndrome variant 2 among the 6 individual factor syndrome variant case definitions (table 4). The rate of the overall case definition was significantly greater in the deployed than the deployable nondeployed in all groups studied except those serving as Air Force aircraft maintenance (table 4).

Association with Functional Status

Even though the factor analysis detected nonrandom symptom patterns without regard to severity of illness, deployed veterans meeting the overall case definition and its component syndrome variant case definitions had significantly lower mean functional status, adjusted for age, sex and race/ethnicity and the analysis weights, than the nonsyndromic veterans on the SF-12 physical component and mental component summary scales [46] (table 5). The effect sizes (difference from the nonsyndromic group divided by the standard deviation of that group) of all groups meeting the case definition ranged between 1.0 and 2.0, indicating very large losses of health-related quality of life in both physical and mental functioning [47] (table 5).

Discussion

The findings from this national survey provide evidence supporting the usefulness of the original factor analysis-derived case definition with 3 primary variants

Table 3. Goodness-of-fit statistics for structural equation model of Gulf War illness with 3 first-order factors (syndrome variants) and a second-order factor (overall Gulf War illness)¹, by study and sample within study

Study and sample within study	Sample size	Goodness-of-fit statistics			
		SRMR	RMSEA	CFI	TLI
Criteria for a good fit [49]		≤0.080	≤0.060	≥0.950	≥0.950
Deployed US Navy Seabees battalion (developmental sample) [8]	249	0.043	0.023	0.992	0.988
Deployed US Army veterans (first validation sample) [8]	335	0.043	0.044	0.975	0.964
USMHS ²					
Random half 1	3,408	0.054	0.018	0.968	0.954
Random half 2	3,408	0.048	0.017	0.972	0.967
Both halves combined	6,816	0.048	0.017	0.970	0.958
Forced equal loadings across both halves	6,816	0.054	0.015	0.972	0.967

SRMR = Standardized root mean-square residual, an absolute fit index, analogous to R^2 for a linear model, and the most sensitive to misspecification of factor covariances or latent structures; the remaining 3 fit indexes are most sensitive to misspecification of factor loadings [49]. Hu and Bentler [49] reported that the combination of SRMR >0.09 and root mean-square error of approximation >0.06 for rejection results in the least sum of type I and type II model rejection errors; RMSEA = root mean-square error of approximation, an absolute fit index that adjusts fit by the number of model parameters estimated to prevent large complex mod-

el structures from inflating the fit [49]; CFI = comparative fit index, a type 3 incremental fit index that estimates the improvement in fit over a baseline null model where all measured variables are uncorrelated [49]; TLI= Tucker-Lewis index (also Bentler-Bonett nonnormed fit index), a type 2 incremental fit index [49].

¹ Corresponds to model 3 developed in the earlier confirmatory factor analysis study [8], pictured in figure 3. ² All results are population estimates adjusted to correct for unequal selection probabilities and minimize bias from nonlocation and nonparticipation by application of the survey analysis weights.

Fig. 3. A structural equation model of Gulf War illness with 3 first-order factors (syndrome variants), each with 4 symptom scales loading on it, and a second-order factor (overall Gulf War illness). The model was developed by a 1995 exploratory factor analysis in 249 members of a US Navy construction battalion [5], validated by a later confirmatory factor analysis in 335 primarily US Army veterans from a VA medical center [8], and in the current study validated with population estimates from the 2007–2009 USMHS, excluding military personnel selected for special studies (see fig. 2). The model shown in the figure corresponds to model 3 validated in the earlier factor analysis study [8].

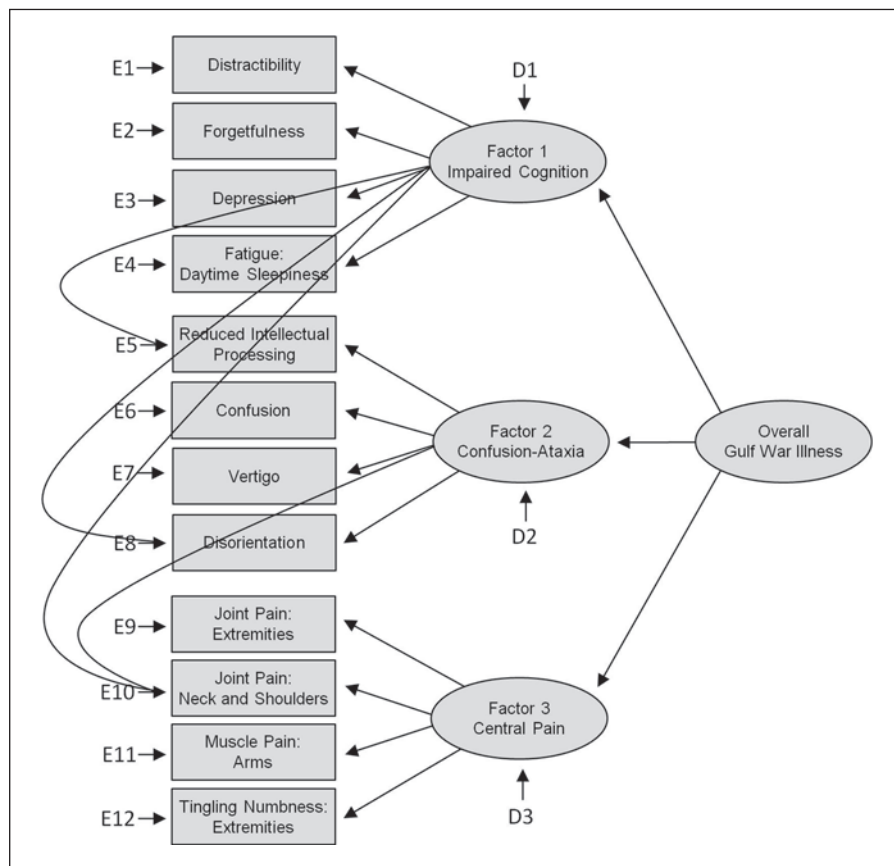


Table 4. Adjusted odds ratio (aOR) for meeting the case definition in deployed versus deployable nondeployed populations, by case definition and by demographic and military characteristics

Case definition and domain	Percent meeting case definition		aOR	95% CI	Tests of the prestatd hypothesis
	deployable nondeployed	deployed			
<i>Overall factor case definition</i>	3.98	13.59	3.87	2.61–5.74	0.001
<i>Factor syndrome variant case definitions</i>					
Syndrome variant 1: impaired cognition	0.59	1.76	3.33	1.10–10.10	0.033
Syndrome variant 2: confusion-ataxia	1.12	6.10	5.11	2.43–10.75	0.001
Syndrome variant 3: central neuropathic pain	1.22	4.58	4.25	2.33–7.74	0.001
Syndrome variant 4: phobia-apraxia	1.41	5.31	3.44	1.75–6.76	0.001
Syndrome variant 5: fever-adenopathy	0.98	1.78	2.06	1.02–4.12	0.042
Syndrome variant 6: weakness-incontinence	0.70	1.20	1.48	0.55–3.94	0.437
<i>Overall factor case definition by domain</i>					
<i>Age</i>					
<49 years	2.72	13.68	5.50	3.09–9.81	0.001
≥49 years	5.88	13.37	2.55	1.49–4.35	0.001
<i>Gender</i>					
Male	3.23	13.40	4.27	2.65–6.90	0.001
Female	7.61	15.37	2.30	1.24–4.27	0.008
<i>Race/ethnicity</i>					
Non-Hispanic White	3.33	9.77	3.56	2.25–5.63	0.001
Black/other	6.78	20.95	4.32	2.10–8.89	0.001
<i>Component</i>					
Active duty	3.23	13.51	4.39	2.43–7.92	0.001
Reserve/guard	4.82	13.91	3.53	2.06–6.06	0.001
<i>Occupation</i>					
Air flight crew	0.35	1.68	10.04	1.87–53.81	0.007
Aircraft maintenance	3.63	6.37	1.56	0.34–7.21	0.568
Army special forces	2.08	17.92	10.15	1.19–86.60	0.034

CI = Confidence interval. All results are population estimates adjusted to correct for unequal selection probabilities and minimize bias from nonlocation and nonparticipation by application of the survey analysis weights. Odds ratios adjusted for age, gender and race/ethnicity using the logistic regression procedure (proc logist) in SUDAAN®. The overall factor case definition is defined

as satisfying any of the 6 factor syndrome variant case definitions. Factor syndrome variants 1–3 are considered the primary syndrome variants because factor syndrome variants 4–6 overlap strongly with factor syndrome variant 2, which is associated with the greatest reduction in functional status and the severest neuropsychological and neuroimaging abnormalities [7, 10–17].

as a case definition for research on Gulf War illness. That the case definition was originally developed by factor analysis of symptoms in a single battalion was shown to fit well a validation sample drawn from ill Gulf War veterans attending a VA clinic, and was associated with objective tests of illness in similarly small pilot studies suggested its usefulness but left open the questions of its external validity. Our present findings address this question in 3 ways.

First, they show that the structural equation model of the case definition fits well the symptom data collected in

our survey of a stratified random sample of the Gulf War era US veteran population. The survey incorporated state-of-the-art survey techniques to ensure a representative sample of Gulf War era veterans from whom to obtain reports of symptoms, exposures and family effects. The distribution of the factor case definition throughout the target population was demonstrated by showing a good fit of the complex syndromic structure to both random halves of the sample with confirmatory factor analysis by structural equation modeling. The quantitative criteria used to indicate a good fit are evidence-based

Table 5. Health-related quality of life measured by the MOS SF-12 physical and mental component summary scores in deployed Gulf War veterans, by the overall factor case definition and its component syndrome variant case definitions

Clinical groups defined by the case definition	SF-12 physical component summary score		SF-12 mental component summary score	
	mean	effect size	mean	effect size
Does not meet the overall case definition	47.5 (0.3)	reference	54.2 (0.9)	reference
Meets the overall factor case definition	34.7 (1.0)	1.3	39.5 (0.9)	1.5
Meets factor syndrome variant case definitions				
Syndrome variant 1: impaired cognition	35.1 (1.6)	1.2	36.9 (1.2)	1.7
Syndrome variant 2: confusion-ataxia	34.9 (1.9)	1.3	33.8 (1.1)	2.0
Syndrome variant 3: central neuropathic pain	32.7 (1.7)	1.5	42.4 (2.3)	1.2
Syndrome variant 4: phobia-apraxia	32.8 (1.3)	1.5	35.7 (1.3)	1.9
Syndrome variant 5: fever-adenopathy	37.6 (1.5)	1.0	43.0 (1.7)	1.1
Syndrome variant 6: weakness-incontinence	31.4 (1.7)	1.6	35.9 (1.9)	1.8

Results are means, with standard error of the mean in parentheses. The two component scores are T scores with a US reference population mean of approximately 50 and a standard deviation of 10. All results are population estimates adjusted to correct for unequal selection probabilities and minimize bias from nonlocation and nonparticipation by application of the survey analysis weights. Least squares mean and standard error of the mean were calculated with the linear modeling procedure of SUDAAN (proc

regress) adjusting for age, gender and race/ethnicity, allowing for the complex stratified sampling design, and weighting by the analysis weights. Effect size is the difference in means of the nonill (not meeting the case definition) and ill groups divided by the standard deviation in the nonill group; an effect size of 0.2 is considered clinically a small effect, one of 0.5 a moderate effect and of 0.8 a large effect [47].

thresholds that maximize the rejection of poorly fitting models and the acceptance of good fitting ones [48–50], particularly the combination of the standardized root mean-square residual and root mean-square error of approximation fit indices [49]. Analysis weights incorporating corrections for unequal selection probabilities among strata and for nonlocation and nonparticipation were applied to the structural equation model analyses to facilitate the unbiased estimation of inferential population parameters from the survey sample.

Second, we found the prevalence of veterans meeting the overall case definition to be low in the deployable nondeployed Gulf War era veteran population and approximately 4-fold higher in the deployed force, as would be expected from an illness caused by exposures in the war theater. In making this comparison, we addressed the well-known selection bias from the fact that only the healthiest soldiers are deployed to a war zone (‘healthy-warrior effect’) [20–23] by omitting from analysis veterans who were nondeployed because of a definable health problem, leaving the deployable nondeployed as the comparable referent group.

Third, the case definition identified groups of deployed Gulf War veterans with greatly reduced functional capacity typical of other serious chronic diseases, as

measured by the MOS SF-12 functional status scales [46, 51]. The effect size of the difference between the groups meeting the case definition and those not meeting it was large, 1.0–2.0. According to Cohen’s rule of thumb, an effect size of 0.2 is considered a small effect, one of 0.5 a moderate effect and of 0.8 a large effect [47]. The SF-36 is the most widely used multi-item instrument for assessing health-related quality of life, and the SF-12 is a short version containing a 12-item subset that strongly predicts the results of the full SF-36 and that is better suited for large surveys. Application of the full SF-36 in a small sample nested in an epidemiological survey of a single battalion previously found that the ill Gulf War veterans meeting the factor syndrome case definitions had reductions in scores on the physical and mental component summary scores that were both statistically and medically significant [17]. The present study confirms that finding in a representative population sample of deployed Gulf War veterans.

Basing a case definition on empirically derived combinations of symptoms shown to occur uncommonly in the deployable nondeployed population and far more commonly in the deployed population and derived from a relatively high cut point (1.5 standard deviations) on the syndrome variant factor scales, maximizes its specificity,

which, by misclassifying few noncases as cases, is optimal for research on etiology, pathogenesis and treatment. However, high specificity is often accompanied by lower sensitivity, in this instance by excluding cases just below the syndrome variant factor scale cut point and ones with rare or unique symptom patterns. Consequently, this research case definition may not prove optimal for eventual clinical screening, treatment and decision-making on service connection of disabilities, where objective biological measures may be preferable.

To identify such objective measures, we have selected two sequential nested case-control subsamples from the participants in this national survey for more efficient application of expensive clinical research techniques, pursuing hypotheses developed in prior research on smaller convenience samples [7, 10–17]. The first subsample is comprised of all subjects meeting the overall factor case definition or the modified Kansas [18] or CDC [38] case definitions (cases) and a random subsample of those not meeting any case definition (controls). These subjects were contacted and asked to provide a blood sample for banking of serum, plasma, DNA and RNA, primarily for testing gene-environment interactions relevant to inferring the original causes and pathogenetic mechanisms of the illness [25, 52, 53]. The second is a smaller random subsample of the cases and controls who provided a blood sample; they have undergone extensive clinical testing in-

cluding multimodal brain imaging, high-resolution electroencephalography and other biomarker measurements [15, 16]. The results of these clinical studies, to be described in future articles, should eventually form the basis for an objective, optimally sensitive and specific disease definition for medical use developed in statistical samples of the target population of Gulf War veterans.

Acknowledgements

This work was supported by US Army Medical Research and Materiel Command grants DAMD17-97-2-0725 and DAMD17-01-1-0741 and contract VA549-P-0027 (Robert W. Haley, PI) administered by the US Department of Veterans Affairs Medical Center, Dallas, Tex., USA. The content does not necessarily reflect the position or the policy of the Federal government or the sponsoring agencies, and no official endorsement should be inferred. A large research team of survey specialists from RTI International contributed importantly to the design and performed the field work for the national CATI survey of Gulf War era veterans. Research leaders included Kathleen A. Considine, Vincent G. Iannacchione, Jill A. Dever, Christopher P. Carson, Heather Best, Carla Bann, Darryl Creel, Barbara Alexander, Amanda Lewis-Evans, Lily Trofimovich, Kirk Pate, Anne Kenyon, Jeremy Morton, Craig Hill and Robert E. Mason. E. William Byrd Jr., Michael E. Murray, Helen Koo and a team of RTI staff contributed to the design of reproductive and child development issues of the CATI questionnaire. Rick Thompson, Eric Cordell and Jennifer Escobar provided program management at UT Southwestern.

References

- 1 Research Advisory Committee on Gulf War Veterans' Illnesses: Gulf War illness and the health of Gulf War veterans. Washington, Department of Veterans Affairs, 2008. <http://www1.va.gov/rac-gwvi/>.
- 2 DeFraités RF, Wanat RR, Norwood AE, William S, Cowan D, Callahan T: Investigation of a suspected outbreak of an unknown disease among veterans of Operation Desert Shield/Storm, 123rd Army Reserve Command, Fort Benjamin Harrison, Indiana, April 1992. Washington, Walter Reed Army Institute of Research, 1992. <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA301076&Location=U2&doc=GetTRDoc.pdf>.
- 3 Berg SW: Post-Persian Gulf medical findings in military reservists. Presentation to the NIH Technology Assessment Conference on the Persian Gulf experience and health, 27–29 April 1994, Bethesda, MD, 1994. http://www.gulfink.osd.mil/seabee/med_270494.pdf.
- 4 Joseph SC: A comprehensive clinical evaluation of 20,000 Persian Gulf War veterans: Comprehensive Clinical Evaluation Program evaluation team. *Milit Med* 1997;162:149–155.
- 5 Haley RW, Kurt TL, Hom J: Is there a Gulf War syndrome? Searching for syndromes by factor analysis of symptoms. *JAMA* 1997;277:215–222.
- 6 Haley RW, Kurt TL: Self-reported exposure to neurotoxic chemical combinations in the Gulf War: a cross-sectional epidemiologic study. *JAMA* 1997;277:231–237.
- 7 Haley RW, Spence JS, Carmack PS, Gunst RF, Schucany WR, Petty F, Devous MD Sr, Bonte FJ, Trivedi MH: Abnormal brain response to cholinergic challenge in chronic encephalopathy from the 1991 Gulf War. *Psych Res Neuroimag* 2009;171:207–220.
- 8 Haley RW, Luk GD, Petty F: Use of structural equation modeling to test the construct validity of a case definition of Gulf War syndrome: invariance over developmental and validation samples, service branches and publicity. *Psychiatr Res* 2001;102:175–200.
- 9 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 4 (DSM-IV). Washington, American Psychiatric Association, 1994.
- 10 Haley RW, Hom J, Roland PS, Bryan WW, Van Ness PC, Bonte FJ, Devous MD Sr, Mathews D, Fleckenstein JL, Wians FH Jr, Wolfe GI, Kurt TL: Evaluation of neurologic function in Gulf War veterans: a blinded case-control study. *JAMA* 1997;277:223–230.
- 11 Hom J, Haley RW, Kurt TL: Neuropsychological correlates of Gulf War syndrome. *Arch Clin Neuropsychol* 1997;12:531–544.
- 12 Roland PS, Haley RW, Yellin W, Owens K: Vestibular dysfunction in Gulf War syndrome. *Otolaryngol Head Neck Surg* 2000;122:319–329.
- 13 Haley RW, Vongpatanasin W, Wolfe GI, Bryan WW, Armitage R, Hoffmann RF, Petty F, Callahan TS, Charuvastra E, Shell WE, Marshall WW, Victor RG: Blunted circadian variation in autonomic regulation of sinus node function in veterans with Gulf War syndrome. *Am J Med* 2004;117:469–478.

- 14 Haley RW, Marshall WW, McDonald GG, Daugherty M, Petty F, Fleckenstein JL: Brain abnormalities in Gulf War syndrome: evaluation by ¹H magnetic resonance spectroscopy. *Radiology* 2000;215:807–817.
- 15 Briggs RW, Cheshkov S, Lu H, Li X, McColl RW, Buhner D, Ferree T, Haley RW: Objective brain abnormalities by 3T MRI and EEG in Gulf War illness. 49th Annual Meeting of the Society of Toxicology, Salt Lake City, March 2010, Biomarkers Session, abstract No 2293.
- 16 Ringe W, Briggs RW, Gopinath K, Kraut M, Rypma B, Odegard T, Bartlett J, Crosson B, Hart J, Haley RW: Functional neuroimaging shows abnormalities in brain function underlying symptoms of Gulf War illness. 49th Annual Meeting of the Society of Toxicology, Salt Lake City, March 2010, Biomarkers Session, abstract No 2294.
- 17 Haley RW, Maddrey AM, Gershenfeld HK: Severely reduced functional status in veterans fitting a case definition of Gulf War syndrome. *Am J Publ Health* 2002;92:46–47.
- 18 Steele L: Prevalence and patterns of Gulf War illness in Kansas veterans: association of symptoms with characteristics of person, place, and time of military service. *Am J Epidemiol* 2000;152:992–1002.
- 19 Poblete PP, Araneta MRG, Sato PA, Hiliopoulos KM, Kamens DR, Morn CB, Zau AC, Gray GC: National study on reproductive outcomes: a reliability study of self-administered survey vs telephone interview. Conference of Federally Sponsored Gulf War Veterans' Illnesses Research, Washington, June 1998, p 46.
- 20 Haley RW: Point: bias from the 'healthy-warrior effect' and unequal follow-up in three government studies of health effects of the Gulf War. *Am J Epidemiol* 1998;148:315–323.
- 21 Larson GE, Highfill-McRoy RM, Booth-Kewley S: Psychiatric diagnoses in historic and contemporary military cohorts: combat deployment and the healthy warrior effect. *Am J Epidemiol* 2008;167:1269–1276.
- 22 McLaughlin R, Nielsen L, Waller M: An evaluation of the effect of military service on mortality: quantifying the healthy soldier effect. *Ann Epidemiol* 2008;18:928–936.
- 23 Wilson J, Jones M, Fear NT, Hull L, Hotopf M, Wessely S, Rona RJ: Is previous psychological health associated with the likelihood of Iraq War deployment? An investigation of the 'healthy warrior effect'. *Am J Epidemiol* 2009;169:1362–1369.
- 24 Defense Science Board: Report of the Defense Science Board Task Force on Persian Gulf War health effects. Washington, Office of the Under Secretary of Defense for Acquisition and Technology, 1994. <http://www.gulfink.osd.mil/dsbrpt/>.
- 25 Haley RW, Billecke S, La Du BN: Association of low PON1 type Q (type A) arylesterase activity with neurologic symptom complexes in Gulf War veterans. *Toxicol Appl Pharmacol* 1999;157:227–233.
- 26 Kang HK, Mahan CM, Lee KY, Murphy FM, Simmens SJ, Young HA, Levine PH: Evidence for a deployment-related Gulf War syndrome by factor analysis. *Arch Environ Health* 2002;57:61–68.
- 27 Cherry N, Creed F, Silman A, Dunn G, Baxter D, Smedley J, Taylor S, Macfarlane GJ: Health and exposures of United Kingdom Gulf war veterans. I. The pattern and extent of ill health. *Occup Environ Med* 2001;58:291–298.
- 28 Blanchard MS, Eisen SA, Alpern R, Karlinksky J, Toomey R, Reda DJ, Murphy FM, Jackson LW, Kang HK: Chronic multisymptom illness complex in Gulf War I veterans 10 years later. *Am J Epidemiol* 2006;163:66–75.
- 29 Horner RD, Kamins KG, Feussner JR, Grambow SD, Hoff-Lindquist J, Mitsumoto H, Pascuzzi R, Spencer P, Tim R, Howard D, Smith TC, Ryan MA, Coffman CJ, Kasarskis EJ: Occurrence of amyotrophic lateral sclerosis among Gulf War veterans. *Neurology* 2003;61:742–749.
- 30 Institute of Medicine: Gulf War and Health. Washington, National Academies Press, 2004, vol 1: Depleted Uranium, Pyridostigmine Bromide, Sarin, Vaccines; vol 2: Insecticides and Solvents – Updated Literature Review on Sarin; vol 3: Fuels, Combustion Products, and Propellants. <http://www.iom.edu/report.asp?id=24236>.
- 31 Araneta MR, Moore CA, Olney RS, Edmonds LD, Karcher JA, McDonough C, Hiliopoulos KM, Schlagen KM, Gray GC: Goldenhar syndrome among infants born in military hospitals to Gulf War veterans. *Teratology* 1997;56:244–251.
- 32 Iowa Persian Gulf Study Group: Self-reported illness and health status among Gulf War veterans: a population-based study. The Iowa Persian Gulf Study Group. *JAMA* 1997;277:238–245.
- 33 Nagelkerke NJD: A note on a general definition of the coefficient of determination. *Biometrika* 1991;78:691–692.
- 34 Dever JA, Mason RE: DMDC Sampling Planning Tool (Version 2.1). Arlington, Defense Manpower Data Center, 2003.
- 35 Kuhn WW, Tucker AW: Nonlinear Programming. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, University of California Press, 1951, pp 481–492.
- 36 Mason RE, Wheelless SC, George BJ, Dever JA, Riemer RA, Elig TW: Sample Allocation for the Status of the Armed Forces Surveys. Proceedings of the Section on Survey Research. Alexandria, American Statistical Association, 1995, pp 769–774.
- 37 Hillier FS, Lieberman GJ: Operations Research, ed 2. San Francisco, Holden-Day, 1974.
- 38 Fukuda K, Nisenbaum R, Stewart G, Thompson WW, Robin L, Washko RM, Noah DL, Barrett DH, Randall B, Herwaldt BL, Mawle AC, Reeves WC: Chronic multisymptom illness affecting Air Force veterans of the Gulf War. *JAMA* 1998;280:981–988.
- 39 American Association for Public Opinion Research: Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, ed 5. Lenexa, AAPOR, 2008.
- 40 Groves R, Couper M: Nonresponse in Household Interview Surveys. New York, Wiley & Sons, 1998.
- 41 Iannacchione V: Sequential weight adjustments for location and cooperation propensity for the 1995 National Survey of Family Growth. *J Offic Statist* 2010;19:31–43.
- 42 Breiman L, Friedman J, Stone C, Olshen R: Classification and Regression Trees. New York, CRC Press, 1984.
- 43 Muthen LK, Muthen BO: Mplus User's Guide, ed 5. Los Angeles, Muthen & Muthen, 2007.
- 44 Ware J Jr, Kosinski M, Keller SD: SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, New England Medical Center, 1995.
- 45 SUDAAN Language Manual, Release 10.0. Research Triangle Park, Research Triangle Institute, 2008.
- 46 Ware J Jr, Kosinski M, Keller SD: A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–233.
- 47 Hays RD, Farivar SS, Liu H: Approaches and recommendations for estimating minimally important differences for health-related quality of life measures. *COPD J Chron Obstruct Pulm Dis* 2005;2:63–67.
- 48 Hu L-T, Bentler PM: Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods* 1998;3:424–453.
- 49 Hu L-T, Bentler PM: Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equat Modeling* 1999;6:1–55.
- 50 Byrne BM: Structural Equation Modeling with EQS and EQS/Windows: Basic Concepts, Applications, and Programming. Thousand Oaks, Sage Publications, 1994.
- 51 Ware JE, Snow KK, Kosinski M, Gandek B: SF-36 Health Survey Manual and Interpretation Guide. Boston, Health Institute, 1997.
- 52 Haley RW, Kramer GL, Xiao J, Teiber JF: Nerve agent exposure associated with Gulf War encephalopathy through gene-environment interaction with the Q192R polymorphism of paraoxonase 1 (PON1). 49th Annual Meeting of the Society of Toxicology, Salt Lake City, March 2010, Neurotoxicity and Neurodegenerative Disease Session, abstract No 2242.
- 53 Hays RD, Sherbourne CD, Spritzer KL, Dixon WJ: A microcomputer program (sf36.exe) that generates SAS code for scoring the SF-36 health survey. Proceedings of the 22nd Annual SAS Users Group International Conference, 1996, pp 1128–1132. <http://gim.med.ucla.edu/FacultyPages/Hays/util.htm>.

Downloaded from <http://karger.com/heap/article-pdf/37/2/129/3153937/000331478.pdf> by guest on 07 December 2023