Epidemiologic Evidence of Health Effects from Long-Distance Transit of Chemical Weapons Fallout from Bombing Early in the 1991 Persian Gulf War

Robert W. Haley\textsuperscript{a}    James J. Tuite\textsuperscript{b}

\textsuperscript{a}Division of Epidemiology, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Tex., and \textsuperscript{b}International Security and Risk Management Consultant (former National Security Advisor to the President pro tempore, US Senate), Gretna, Va., USA

Abstract

\textbf{Background:} Military intelligence data published in a companion paper explain how chemical fallout from US and Coalition bombing of Iraqi chemical weapons facilities early in the air campaign transited long distance, triggering nerve agent alarms and exposing US troops. We report the findings of a population-based survey designed to test competing hypotheses on the impact on chronic Gulf War illness of nerve agent from early-war bombing versus post-war demolition. \textbf{Methods:} The US Military Health Survey performed computer-assisted telephone interviews of a stratified random sample of Gulf War-era veterans (n = 8,020). Early-war exposure was measured by having heard nerve agent alarms and post-war exposure, by the computer-generated plume from the Khamisiyah demolition. Gulf War illness was measured by two widely published case definitions. \textbf{Results:} The OR (95\% CI) for the association of alarms with the Factor case definition was 4.13 (95\% CI 2.51–6.80) compared with 1.21 (95\% CI 0.86–1.69) for the Khamisiyah plume. There was a dose-related trend for the number of alarms ($p_{\text{trend}} < 0.001$) but not for the number of days in the Khamisiyah plume ($p_{\text{trend}} = 0.17$). \textbf{Conclusions:} Exposure to low-level sarin nerve agent in fallout from bombing early in the air campaign contributed more to chronic illness than post-war demolition.

Introduction

During or immediately after the 1991 Persian Gulf War, a large proportion (perhaps 25\% [1]) of the approximately 700,000 deployed US military personnel developed an unusual illness manifested by fatigue, fever and night sweats, memory and concentration problems, pathogen-free diarrhea, sexual dysfunction, chronic body pain and other symptoms compatible with autonomic nervous system dysfunction [2] and damage to the brain’s cholinergic system [2–4], referred to as Gulf War illness. The illness was so florid and widespread that military medical authorities initially suspected an epidemic of occult leishmaniasis [5], eventually not confirmed. When the symptoms persisted, many of the affected military personnel, recalling the repetitive sounding of...
chemical nerve agent alarms in their camps requiring them to don impervious protective suits and gas masks early in the air war, suspected that they were suffering the chronic effects of exposure to low-level chemical nerve agent. An initial investigation in 1994 by a US Senate subcommittee with access to classified operational information documented the widespread sounding of chemical nerve agent alarms throughout US and Coalition troop positions and the detection of low-level sarin nerve agent and blistering agent by chemical experts beginning around the third day of the bombing campaign (Operation Desert Shield, 17 January–23 February 1991), simultaneous with US and Coalition bombing of nerve agent stockpiles deep in Iraq [6]. At the time, chemical weapon experts believed the detections were from fallout from the bombing [7]. A pivotal 1994 report by the Defense Science Board seriously considered this explanation but ultimately rejected it, concluding that the bombed chemical weapons production and storage facilities near Baghdad were too far from the troop positions and surface wind directions were incompatible, leaving the widespread nerve agent alarms and detections early in the air campaign unexplained [7, 8]. Subsequent investigations turned to other war theater exposures such as pesticides, pyridostigmine bromide anti-nerve-agent medication, multiple immunizations, depleted uranium, and combat stress [9]. In a companion paper, Tuite and Haley [10] provide the first published analysis of the evidence explaining how fallout from the air bombing early in the air campaign probably did transit the long distances required to set off nerve agent alarms in US troop positions.

In 1996 a second possible source of low-level chemical weapon exposure was discovered in retrospect [11]. In October 1991 a United Nations Special Commission inspection team visited a large Iraqi storage site at Khamisiyah (see location in fig. 1 in the companion paper [10]). Just south of the main complex of over 100 storage bunkers containing mostly conventional munitions, they reported finding a pit containing 297 122-mm rockets, some damaged but mostly intact, samples from which tested positive for sarin and cyclosarin. In 1996 an investigation in conjunction with the Presidential Advisory Committee on Gulf War Veterans’ Illnesses discovered that elements of the 37th Engineer Battalion had destroyed some munitions in the Khamisiyah pit on 10 March 1991, 10 days after the end of the ground war (Operation Desert Storm, 24–28 February 1991). At the time of the demolition activities, inventory of a sample of the munitions in the pit did not identify Iraqi markings of chemical weapons; consequently, no one was aware of the presence of nerve agent, and post demolition monitoring for ambient nerve agent release by the demolitions was negative. No personnel reported acute symptoms or early effects (e.g., miosis) of nerve agent exposure following the demolitions.

After a series of progressively escalating estimates of the number of troops who might have been exposed if low-level nerve agent had been released from the demolitions aroused public concern [11, 12], the US Department of Defense and the Central Intelligence Agency performed munitions experiments and commissioned a series of computer simulation models to estimate the geographical dispersion pattern of the plume resulting from the Khamisiyah munitions demolition on 10 March 1991 and the subsequent 2–3 days. In September 1997 a composite of 5 simulation models showing different plume patterns but all generally drifting southwesterly from Khamisiyah was released (the ‘1997 plume model’). A sixth computer simulation model, showing a radically different plume pattern drifting to the southeast and then northeast, was excluded [11]. The geographical information system (GIS) coordinates of the 1997 plume model were then overlaid on the GIS coordinates of the positions of US military units in the theater on those days, identifying 98,910 military personnel who might have been exposed to low-level nerve agent [13]. These individuals received written notification of possible exposure. After modification of important base assumptions of the computer simulation modeling and updates to the unit locations database, a revised plume distribution was released in 2000 (the ‘2000 plume model’), this one identifying 101,752 potentially exposed personnel (approximately 65,500 in common with the 1997 model), and correction letters were sent [14]. Both the 1997 and 2000 plume models were subsequently used to identify individuals with low-level nerve agent exposure in various epidemiologic studies of Gulf War illness. In 2004 the Government Accountability Office reported their evaluation of the 1997 and 2000 computer simulation models [11]. Citing serious inaccuracies in the source terms (e.g., quantity and purity of nerve agent), underestimation of explosion plume heights, unrepresentative conditions of field tests, and wide divergence of plume patterns from the various computer plume models, the report concluded that the models do not reliably indicate which troops were exposed to nerve agent, if any was released from the demolitions, and that the statistical power of epidemiologic studies using these plume models would be severely biased toward the null hypothesis by large exposure misclassification [11].
The complex issues in understanding and measuring nerve agent release and exposure have prevented a fair epidemiologic test of the etiologic role of nerve agent in Gulf War illness. This paper reports the results of a study in a large population-representative sample of Gulf War-era veterans designed to test the two competing hypotheses on the source of nerve agent exposure in the 1991 Gulf War against the two most widely published case definitions of Gulf War illness. It represents nerve agent exposure from fallout early in the air war by veterans’ reports of chemical alarms sounding and exposure from the postwar demolition by the 2000 Khamisiyah plume model.

Methods

National Survey of a Population-Representative Sample of Gulf War-Era Veterans

The US Military Health Survey (USMHS), conducted by Research Triangle Institute (Research Triangle Park, N.C.) from 2007 to 2010, involved a computer-assisted telephone interview survey of a population-representative sample of Gulf War-era veterans. The survey objectives, sampling design, allocation, selection of the sample, questionnaire content, data collection and quality control methods, statistical power, methods for constructing sampling weights, and initial findings were published previously in Neuroepidemiology [15]. Briefly, the database of all personnel who were in the US Armed Forces as of August 2, 1990, maintained by the Defense Manpower Data Center (Seaside, Calif., USA), was stratified by age, sex, race, officer versus enlistment rank, military component (active duty vs. activated Reserve/National Guard or recalled retired), special military occupations (air flight crews, aircraft maintenance, army special forces, or other), deployment versus nondeployed to the Kuwaiti Theater of Operations during the conflict period, and in theater location on 20 January 1991. Samples of veterans were selected randomly from each of the strata with unequal selection probabilities designed to provide the most powerful test of the study hypotheses. In-theater location on 20 January 1991 was included in the stratification scheme to ensure optimal representation of troop populations exposed and not exposed to fallout from bombing early in the air campaign. In total, 8,020 Gulf War-era veterans were enrolled and interviewed with questions measuring a broad array of symptoms, wartime exposures and family issues. A sampling weight was constructed for each participant combining a survey design weight (the inverse of the probability of selection into the sample), a nonlocation correction weight derived by a logistic regression analysis to correct for the sample members who could not be located, and a nonparticipation correction weight derived similarly to correct for those who were located but declined to participate [15]. Application of the sampling weights in computing the hypothesis tests removes any selection bias introduced by the unequal selection probabilities and minimizes selection bias from inability to locate or obtain participation from sample nonparticipants, allowing the analytic results to reflect the true population effects.

Case Definitions of Gulf War Illness

Two formal case definitions of Gulf War illness have been used commonly in the scientific literature. The first, known as the Factor case definition [16], was developed in 1995 in 249 members of a Naval Reserve construction battalion of Gulf War veterans by principal components hierarchical factor analysis of symptoms typically endorsed by Gulf War veterans, resulting in 6 orthogonal syndrome factor scales, each trichotomized into the syndromic (≥1.5 standard deviations), the subsyndromic (1.0 to <1.5 standard deviations) and the nonsyndromic (<1.0 standard deviations) level. Subjects satisfied the Overall Factor case definition if they scored ≥1.5 standard deviations on any of the 6 syndrome factor scales. The factor structure was validated by confirmatory factor analysis with structural equation modeling in 335 primarily US Army veterans attending a Gulf War illness clinic at a Veterans Affairs medical center in Texas [17] and again in random split halves of the 8,020 randomly selected participants in the USMHS [15].

The second case definition, known as the CDC Chronic Multisymptom Illness (CMI) case definition [18], was empirically specified to provide a simple definition that could be more easily applied in a clinical setting. A case of CDC CMI was defined as having at least 1 chronic symptom (present for at least 6 months) from at least 2 of the following 3 categories: (a) fatigue, (b) mood or cognition (any symptoms of feeling depressed, difficulty remembering or concentrating, feeling moody, feeling anxious, trouble finding words, or difficulty sleeping), and (c) musculoskeletal (any symptoms of joint pain, joint stiffness, or muscle pain). A case was subclassified as ‘severe’ if both of the case-defining symptoms were rated by the subject as severe; otherwise, it was considered ‘mild to moderate’ in severity. Although no formal validation has been performed, the CDC CMI case definition has been used in many subsequent studies, including two large clinical trials of treatments for Gulf War illness [1, 19, 20].

Exposure Measures

The following survey question was included in the USMHS to measure exposure to low-level nerve agent during the conflict period when nerve agent detector alarms were widely switched on: ‘During the time period from August 2, 1990 to July 31, 1991, did the alarms on the chemical warfare detection devices in areas where you were living or working ever go off while you were present there?’ [15]. A follow-up question asked approximately how many times they heard these chemical warfare detection devices go off.

Individual soldiers’ potential exposure to nerve agent in fallout from the postwar demolition of Iraqi ammunition depots at Khamisiyah on 10 March 1991 – the day of the demolition – and the subsequent 2 days was estimated from the system GIS latitude and longitude coordinates of their military unit’s daily location down to the company level, made available by the US Army Center for Unit Records Research (CURRR). Unit location information was known for 78% of the deployed force. GIS analysis determined whether each subject was located within the map coordinates of the revised computer-generated plume of possible fallout exposure released in 2000 by the US Department of Defense and the Central Intelligence Agency and displayed on the Defense Department’s Gulf War-related website [14]. Theater-deployed military personnel were classified on the binary measure of plume exposure as well as by the number of days located in the plume as.
well as by the number of days in the plume (0, 1 or 2–3 days). There were too few personnel exposed for all 3 days to analyze it separately. A subclassification of plume exposure by estimated level of ambient sarin concentrations, used in 5 of the 17 prior epidemiologic studies conducted by government employees (table 1), was determined by the Office of Deployment Health to be classified and not made available for this study.

Statistical Analysis

Multivariable logistic regression analysis of the USMHS database was performed to assess the independent association of low-level nerve agent exposure with the two widely used case definitions of Gulf War illness. The analysis, performed with SUDAAN software [21], allowed for the complex survey sampling design and applied sampling weights to project findings from the sample to the full Gulf War-era US military population [15] for estimating the numbers of personnel exposed when nerve agent alarms sounded during the conflict period or from putative exposure to the computer-modeled plume from the postconflict demolition of the Khamisiyah ammunition depot. Population relative risks were estimated by a logistic regression analysis comparing the exposure of the deployed veterans in the sample meeting the given case definition with those of all deployed veterans in the sample with no evidence of chronic illness (not meeting any of the case definitions) [16]. Nondeployed personnel were not included in the analysis.

The independent variables included dummy variables representing categories of the number of times nerve agent alarms sounded where the subject was present, the number of days located in the Khamisiyah plume, and the potential confounding variables: age at survey start (<49, ≥49), gender, race/ethnicity (nonblack, black), and rank during the war (enlisted, officer). The 4 categories of the number of times nerve agent alarms sounded were developed in a random half of the USMHS database and confirmed in the second half (n = 3,480 in each subsample). Adjusted population odds ratios and 95% confidence intervals were calculated from dummy variables leaving out exposure category 0 (unexposed) to serve as the referent group. After the final multivariable models testing the effects of nerve agent alarms and the Khamisiyah plume were determined, they were rerun with the addition of independent variables to test for confounding by the computer-modeled plume from the postconflict demolition of the Khamisiyah ammunition depot. Population relative risks were estimated by a logistic regression analysis comparing the exposure of the deployed veterans in the sample meeting the given case definition with those of all deployed veterans in the sample with no evidence of chronic illness (not meeting any of the case definitions) [16]. Nondeployed personnel were not included in the analysis.

The independent variables included dummy variables representing categories of the number of times nerve agent alarms sounded where the subject was present, the number of days located in the Khamisiyah plume, and the potential confounding variables: age at survey start (<49, ≥49), gender, race/ethnicity (nonblack, black), and rank during the war (enlisted, officer). The 4 categories of the number of times nerve agent alarms sounded were developed in a random half of the USMHS database and confirmed in the second half (n = 3,480 in each subsample). Adjusted population odds ratios and 95% confidence intervals were calculated from dummy variables leaving out exposure category 0 (unexposed) to serve as the referent group. After the final multivariable models testing the effects of nerve agent alarms and the Khamisiyah plume were determined, they were rerun with the addition of independent variables to test for confounding by the computer-modeled plume from the postconflict demolition of the Khamisiyah ammunition depot. Population relative risks were estimated by a logistic regression analysis comparing the exposure of the deployed veterans in the sample meeting the given case definition with those of all deployed veterans in the sample with no evidence of chronic illness (not meeting any of the case definitions) [16]. Nondeployed personnel were not included in the analysis.

Tests of the Two Nerve Agent Exposure Hypotheses

In a population-representative multivariable logistic regression model controlling for age, sex, race and military rank, veterans who had been in areas where the chemical nerve agent alarms sounded were 4.13 (95% CI 2.51–6.80) times more likely to have a chronic illness meeting the Overall Factor case definition than those who had not been where alarms sounded (p < 0.001); whereas veterans who had been located in the computer-
Table 1. Published studies on association of health conditions with nerve agent exposure measured by the computer-modeled plume from the Khamisiyah demolition

<table>
<thead>
<tr>
<th>First author (year)</th>
<th>Sample sizes in categories of increasing exposure to nerve agent in the plume</th>
<th>Plume model</th>
<th>Form of the Khamisiyah exposure variable used</th>
<th>Outcome studied</th>
<th>Authors’ conclusion</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray [34] (1999)</td>
<td>177,343/18,952/23,061/6,757 (75,717 uncertain)</td>
<td>1997</td>
<td>3 levels</td>
<td>hospitalization and ICD-9 categories</td>
<td>negative</td>
<td></td>
</tr>
<tr>
<td>Smith [37] (2003)</td>
<td>318,458/99,614</td>
<td>2000</td>
<td>yes</td>
<td>ICD-9 hospital discharge codes</td>
<td>positive</td>
<td>associated only with arrhythmias</td>
</tr>
<tr>
<td>Smith [38] (2004)</td>
<td>474,109/72,413</td>
<td>2000</td>
<td>yes</td>
<td>hospitalization</td>
<td>negative</td>
<td></td>
</tr>
<tr>
<td>Bullman [24] (2005)</td>
<td>224,980/100,487</td>
<td>2000</td>
<td>yes 0, 1, ≥2</td>
<td>ICD-9 cause-of-death codes from the National Death Index</td>
<td>positive</td>
<td>dose-related association of number of days in plume with brain cancer</td>
</tr>
<tr>
<td>Mahan [40] (2005)</td>
<td>3,336/1,898</td>
<td>2000</td>
<td>yes 0, 1, ≥2</td>
<td>general health, clinic, CFS, employment</td>
<td>negative</td>
<td></td>
</tr>
<tr>
<td>Blanchard [42] (2006)</td>
<td>234/92</td>
<td>2000</td>
<td>yes</td>
<td>CDC CMI case definition</td>
<td>positive</td>
<td>significantly associated with severe CMI</td>
</tr>
<tr>
<td>Proctor [44] (2006)</td>
<td>70/47/23</td>
<td>2000</td>
<td>3 levels</td>
<td>neuropsychological tests</td>
<td>positive</td>
<td>associated with psychomotor and visuospatial abilities</td>
</tr>
<tr>
<td>Miranda [46] (2008)</td>
<td>50 ALS vs. 690,000 controls2</td>
<td>2000</td>
<td>yes</td>
<td>ALS</td>
<td>negative</td>
<td>Suggestive association not statistically significant</td>
</tr>
<tr>
<td>Toomey [47] (2009)</td>
<td>1,061 deployed vs. 1,128 nondeployed2</td>
<td>2000</td>
<td>yes</td>
<td>neuropsychological tests</td>
<td>positive</td>
<td>associated with verbal memory</td>
</tr>
<tr>
<td>Barth [25] (2009)</td>
<td>523,496/98,406</td>
<td>2000</td>
<td>0, 1, ≥2</td>
<td>brain cancer, ALS, multiple sclerosis</td>
<td>positive</td>
<td>significant trend with brain cancer</td>
</tr>
<tr>
<td>Chao [48] (2010)</td>
<td>40/40 (7 moderate, 24 high, 9 unknown)</td>
<td>2000</td>
<td>3 levels</td>
<td>MRI VBM-based brain volumes and neuropsychological tests</td>
<td>positive/ negative</td>
<td>binary associated with GM volume; exposure levels not associated</td>
</tr>
</tbody>
</table>

ICD-9 = International Classification of Diseases, 9th ed.; CFS = chronic fatigue syndrome; VBM = volume-based morphometry; ALS = amyotrophic lateral sclerosis; MRI = magnetic resonance imaging; WB = brain white matter; GM = brain gray matter.
1 Unexposed category listed first, followed by either the exposed group for binary measures or categories of increasing exposure concentration for polychotomous measures.
2 Did not report the numbers exposed and nonexposed to the Khamisiyah plume.
3 Did not provide a breakdown sample size of the exposed category by the number of days exposed.
modeled plume from the postconflict demolition of the Khamisiyah ammunition depot on any day were 1.21 (95% CI 0.86–1.69) times more likely to have a chronic illness meeting the Overall Factor case definition than those who had not been located in the plume (p = 0.27). Veterans located where alarms sounded were 2.60 (2.00–3.38) times more likely to have a chronic illness satisfying the CDC CMI case definition (p < 0.001); whereas those located in the Khamisiyah plume were 1.17 (0.93–1.46) times more likely to satisfy the CDC CMI case definition (p = 0.19).

Test for Dose-Response Relationship
The population-estimated relative risk of Gulf War illness, measured by the Overall Factor case definition, increased in a strong dose-related progression with the number of times nerve agent alarms sounded in the subjects’ immediate working or living area during the conflict period up to a relative risk of 8.6 (fig. 2a), but not with the number of days subjects were located in the Khamisiyah plume (fig. 2c). A similar dose-related association with the number of chemical alarms was found with the CDC CMI case definition, although the strength of the association was less (fig. 2b).

Application of the trend test found statistically significant (p trend1 < 0.001) dose-related increases in the risk of each of the case definitions and their subclassifications over the 4 categories of the number of times nerve agent alarms sounded in the subjects’ working or living area while they were present (fig. 2a, b). When the group that reported no chemical alarms (the unexposed group) was omitted, the trend test remained statistically significant for all case definitions (fig. 2a, b) except the syndrome 1 variant of the Factor case definition (p trend2 = 0.93), suggesting that the syndrome 1 variant is less likely than the other syndrome variants to be causally related to sarin exposure [23]. That the mild-to-moderate (p trend2 = 0.15) and severe (p trend2 = 0.079) categories of the CDC CMI case definition remained of borderline significance (fig. 2b) suggests that the drop in significance may have been due more to loss of power from omission of the unexposed group than from confounding [23]. Both the magnitude and the significance of the dose-related associations of the syndrome 4–syndrome 6 variants of the Factor case definition were similar to those of the syndrome 2 variant to which they are closely related [16] (data not shown).

In contrast, the number of days exposed to the computer-modeled Khamisiyah plume was unassociated with the Overall Factor case definition and its three variants (fig. 2c) but showed a statistically significant, though small, dose-related association with the Overall CDC CMI case definition and its severity levels (fig. 2d). Omission of the group with no plume exposure, however, nullified these associations (p trend2 > 0.6), suggesting that they may have been due to confounding [23].

Testing Exposures in Regions of the Intersection of the Two Case Definitions
Cases that met both the Overall Factor and the CDC CMI case definitions (fig. 1) showed a strong, highly significant dose-related association with the number of times chemical alarms sounded (p trend2 < 0.001, which remained significant when the subgroup reporting no chemical alarms sounding was omitted (p trend2 < 0.001;
Moreover, cases that were subsyndromic (maximum factor score 1.0 to <1.5 SD) in the Overall Factor case definition and also satisfied the CDC CMI case definition (fig. 1) showed a significant dose-related association with the number of chemical alarm soundings with or without inclusion of the unexposed group (table 2). Cases meeting the severe CDC CMI case definition but not the Overall Factor case definition or its subsyndromic level (fig. 1) showed a significant dose-related association with the number of alarms, which did not persist after exclusion of the unexposed subgroup; however, the small number of cases in this group and the lower range p value (p = 0.14) suggest a possible weak biological association obscured by low statistical power (table 2). The cases meeting the mild to moderate CDC CMI case definition but not the Overall Factor case definition or its subsyndromic level (fig. 1) showed only a weakly significant association with the number of chemical alarms sounding that did not persist with omission of the unexposed group despite a very large sample size (table 2).

With SUDAAN software, the analyses allow for the complex stratified sampling design and are weighted by the sampling weights so that the results apply to the full Gulf War veteran population. Adjusted odds ratios and 95% confidence intervals were calculated from dummy variables with exposure category 0 (unexposed) left out to serve as the referent group. The first set of trend tests (p trend1 ) was obtained by entering the 4-category alarms and 3-category plume measures as ordered variables in the multivariable logistic regression analysis. The second set (p trend2 ) was obtained similarly except the subjects in the 0 category (unexposed) were omitted to control confounding by differences associated with no exposure versus any exposure [23].

![Graph showing the association between the number of chemical alarms and days exposed to the Khamisiyah plume.](image-url)
posed to the plume, but they did not remain significant ($p_{\text{trend}} > 0.3$) after inclusion of the unexposed group (table 2).

**Tests for Confounding**

None of the findings was appreciably altered by adding to the multivariable models indicators of wearing uniforms treated with pesticides, taking pyridostigmine bromide, and experiencing combat stress measured by the Keane combat exposure scale (data not shown).

**Discussion**

**Implications for Exposure Assessment**

This epidemiologic study found a strong dose-response relationship between the number of times nerve agent alarms sounded in the subject’s immediate area and the risk of having Gulf War illness as defined by either of the widely used case definitions: the Overall Factor and the CDC CMI. In contrast, the number of days exposed to the computer-generated plume from postconflict demolition of the Khamisiyah ammunition depot did not show the same strong dose-related association with the Overall Factor case definition or its 3 syndrome variants, but it did show a slight dose-related association with the Overall CDC CMI which, however, did not survive omission of the unexposed category. Interestingly, the slight association of the plume with the CDC CMI was only found in those who did not meet either the syndromic or subsyndromic levels of the Overall Factor case definition. These findings support an important etiologic role of low-level nerve agent exposure during the conflict period, when nerve agent alarms were continuously monitoring, in causing the relatively characteristic clinical syndromes identified by principal components factor analysis and measurable with the Overall Factor case definition and its syndrome variants. They do not support an important etiologic role of exposure to the postconflict demolitions of the Khamisiyah ammunition depot in causing these same conditions.

These findings support the strongly held view of the chemical warfare experts from several countries who detected chemical nerve and blister agents with sophisticated equipment early in the air campaign and strongly suspected fallout from US and Coalition bombing of Iraqi chemical weapons production and storage facilities deep in Iraq [7] as well as the initial impression of the Defense Science Board reviewing the evidence shortly after the war [8]. The direct evidence for long-distance transit of the warm air mass from the bombings presented in the companion paper [10] finally puts to rest the doubts about long-distance transit that caused the afterwar analyses ultimately to abandon the most strongly suspected explanation for the sudden sounding of tens of thousands of chemical nerve agent alarms and the large epidemic of

<table>
<thead>
<tr>
<th>Regions of the intersection of the Overall Factor and CDC CMI case definitions</th>
<th>Number of cases in sample/estimated number of cases in population</th>
<th>Number of times chemical alarms sounded</th>
<th>Number of days exposed to 2000 Khamisiyah plume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both Overall Factor and CDC CMI</td>
<td>679/60,426</td>
<td>&lt;0.001</td>
<td>0.23</td>
</tr>
<tr>
<td>Both Overall Factor subsyndromic and CDC CMI</td>
<td>450/39,256</td>
<td>&lt;0.001</td>
<td>0.61</td>
</tr>
<tr>
<td>CDC CMI severe only</td>
<td>82/8,489</td>
<td>0.003</td>
<td>0.06</td>
</tr>
<tr>
<td>CDC CMI mild to moderate only</td>
<td>1,072/99,781</td>
<td>0.02</td>
<td>0.07</td>
</tr>
</tbody>
</table>

1 Defined in figure 1.
2 Numbers of cases in the Gulf War veteran population were estimated by applying sampling weights to correct for unequal sampling probabilities among strata and for nonlocation and nonparticipation [15].
3 $p_{\text{trend}}$ was obtained by entering the 4-category alarms (0, 1, 2–20, ≥20 alarms) and 3-category plume (0, 1, 2+ days) measures as ordered variables in the multivariable logistic regression analysis. $p_{\text{trend2}}$ was obtained similarly except the subjects in the 0 category (unexposed) were omitted to control confounding by differences associated with no exposure versus any exposure [23]. The logistic regression analyses were performed with `proc logist` in SUDAAN software controlling for age, sex, race and officer rank, allowing for the complex sampling design, and applying the sampling weights [15].
4 The Overall Factor case definition is ≥1.5 SD on the factor scale; subsyndromic is 1.0 to <1.5 SD [15, 16].
chronic illness that followed. The evidence substantiating the long-distance transit of chemical fallout as the cause of the chemical nerve agent alarms together with the finding of strong dose-related associations of chemical alarms with the chronic symptoms of Gulf War illness, regardless of how defined, bring us full circle back to the original, most apparent explanation, namely, that Gulf War illness is a chronic neurotoxic encephalopathy caused by repetitive exposure to low-level sarin nerve agent in fallout from bombing of chemical weapons production and storage facilities deep in Iraq early in the air campaign.

**Implications for Case Definition**

This study also offers insights into how Gulf War illness is defined. This is important because the difficulty of defining it is the main reason Gulf War illness has remained so controversial, poorly understood and frustrating to veterans and the physicians who care for them. The two case definitions studied here represent opposite ends of the spectrum. The Factor definition was developed with principal components hierarchical factor analysis of a large battery of typical symptoms to identify coherent symptom patterns, each reported by large numbers of ill Gulf War veterans [15, 16]; whereas, the CDC CMI case definition cast a broader net by requiring only one symptom from at least 2 of 3 symptom domains, to provide a definition more simply applied in clinical settings [18]. As predicted from their designs, the Factor definition identifies a smaller nucleus comprising 13.6% of the deployed veteran population that is almost entirely subsumed in the threefold larger body of less specific symptoms comprising 41.7% of the deployed population. Not surprisingly, virtually all of the detectable risk conferred by hearing chemical nerve agent alarms is confined to the Factor definition, and when this nucleus is removed, no robust dose-related association remains with the rest of the CDC CMI definition. What is surprising, however, is that while exposure to the 2000 Khamisiyah plume shows no association with the Factor case definition, it shows a small association with the component of the CDC CMI definition that does not contain the cases meeting the Factor definition. Whereas this dose-related association does not remain after excluding the unexposed category [23], given the large amount of illness misclassification in the CDC CMI definition and the large potential for misclassification of exposure in the computer-generated Khamisiyah plume model [11], one wonders if this small signal might indicate a risk association with some condition other than Gulf War illness hidden in the region of the CDC CMI definition not overlapping the Factor definition. As one possible example, the risk of glial brain cancer has been found in 2 large epidemiologic studies to be significantly associated with the number of days exposed to the 2000 Khamisiyah plume [24, 25].

**Strengths and Weaknesses of This Study**

A strength of this epidemiologic study is its population-representative design and the large sample size (n = 8,020), which was undertaken and designed specifically to validate the Factor case definition and to test its associations with the measures of nerve agent exposure [15]. The survey design effort began in 1997 shortly after publication of the first published evidence linking nerve agent exposure to the Factor case definition [26]. It continued for over 10 years as funding waxed and waned, until the field work was finally conducted from 2007 to 2010. The design and field work were managed by a team of survey specialists of RTI International (Research Triangle Park, N.C., USA), which has had extensive experience in designing and conducting surveys of military populations. The use of the two most widely used case definitions of Gulf War illness provides comparability with much of the scientific literature on the subject. The CDC CMI case definition has been used in many epidemiologic studies, clinical trials of treatment, and studies on the nature of the illness, reviewed by the Research Advisory Committee on Gulf War Veterans’ Illnesses [1]. The Factor case definition has been used most extensively in clinical studies of the illness using neurophysiologic, brain imaging and EEG testing, recently reviewed by lannacchione et al. [15]. Both definitions have been used in neuropsychological and functional capacity studies. Finding strong dose-related associations of nerve agent alarms and weak-to-null associations of the Khamisiyah plume exposure with both case definitions provides assurance that the conclusions are not due to some idiosyncrasy of case definition.

The main potential weakness relates to the reliability of the two measures of sarin exposure used in the analyses. The analysis of possible sarin exposure from the post-conflict demolition of the Khamisiyah ammunition depot was based on computer-modeling of the fallout plume [14] and overlay of the plume coordinates with daily location coordinates of company-level military units to obtain individual subject exposure and thus did not rely on personal recollections. This was particularly important because plume exposure would have been largely unobservable; nerve agent alarms were no longer widely acti-
vated after the end of the conflict; and with many post-conflict demolitions going on, soldiers would not have known whether a witnessed demolition was the actual Khamisiyah event. The Khamisiyah plume model has several weaknesses. (i) A critical evaluation by the US Government Accountability Office concluded that the Khamisiyah modeling process was invalid [11], and yet, with no alternatives researchers continued using it as the measure of nerve agent exposure in government-funded studies (11 of the 17 studies using the modeled exposure measure were published after the US Government Accountability Office report was released; table 1). (ii) A subclassification of the plume into model-predicted levels of ambient sarin concentrations and cumulative sarin exposure – used in 5 of the 17 government-funded studies (table 1) – was classified and unavailable to this study. However, 14 of the 17 past studies of the Khamisiyah plume model used the binary indicator of plume exposure, and 3 used the number of days of plume exposure – both employed in this study – and 7 of these 14 studies showed significant associations with chronic illness (table 1). In this study the associations of the case definitions of Gulf War illness with the binary plume indicator and the number of days of exposure were so weak that further subclassification is unlikely to bring out any strong association. It is possible, however, that other conditions with symptoms unlike Gulf War illness, such as brain cancer [24, 25], might have been caused by exposure to fallout from the Khamisiyah demolition.

The main concern about the survey interview measure of having been in an area where the nerve agent alarms sounded is the possibility of misclassification in recall of such exposure. In contrast to the high likelihood of errors in recall of exposure to the Khamisiyah demolition, the likelihood of misclassifying recall of nerve agent alarms is small for the following reasons: (i) the alarms located in virtually every company-sized camp emitted a sound designed to be unmistakable, (ii) all US military personnel had been extensively trained to recognize it and to take immediate protective action, (iii) after hearing alarms, personnel usually donned protective rubber suits and gas masks, (iv) experiencing a possible nerve agent exposure was generally a frightening event that made a strong, lasting impression, and (v) in repeated surveys of the same large group of Gulf War veterans, reports of having heard alarms were found to have moderately good reliability over time [27]. Furthermore, the script of the computer-assisted interviews followed in the USMHS had the interviewers remind the subjects of the following considerations: (i) the importance of candor and accuracy of responses to make the study valid because untruthful answers are likely to bias toward null findings, and (ii) the subjects’ responses would be protected from all disclosure by an official National Institutes of Health Certificate of Confidentiality, thus removing the expectation of direct secondary gain from biased responses.

The validity of measuring sarin exposure from reports of chemical nerve agent alarms can be further tested by analyzing for evidence of a gene-environment interaction between exposure to chemical alarms and the Q192R polymorphism of the PON1 gene. The PON1 gene and the serum activity of its polymorphic Q and R isoenzyme products are the main determinants of the level of an individual’s resistance to the brain effects of organophosphate pesticides and nerve agents, including sarin [28–30]. The level of serum activity of the Q isoenzyme, the main determinant of susceptibility to sarin, varies widely in the population, and was found in a small study to be associated with Gulf War illness [31, 32]. If the dose-related association between the number of chemical nerve agent alarms reported and Gulf War illness were mostly due to recall bias, one would not expect the alarms measure to show a strong synergistic relationship with the PON1 polymorphism and the serum activity level of the Q isoenzyme. Finding such a gene-environmental interaction would further support the etiologic role of exposure to low-level sarin from bombing early in the air campaign.

Conclusions

The findings of this epidemiologic study, designed specifically to test two competing hypotheses on the source of low-level nerve agent exposure of US Gulf War veterans, support the potential importance to veterans’ health of sarin exposure from fallout produced by bombing early in the air campaign and the relatively less importance of the postconflict Khamisiyah demolition. This is not to deny any role of the Khamisiyah demolitions, which have been implicated by several prior studies, but rather to suggest that the magnitude of postwar chronic illness ascribed to Khamisiyah exposure greatly underestimates the true toll from nerve agent exposure in the 1991 Persian Gulf War. Epidemiologic studies testing only the effects of Khamisiyah exposure are likely to overlook risk to larger groups who were exposed to fallout early but were not in the limited Khamisiyah plume. Where possible, the data from prior epidemiologic studies should be reassessed with multivariable analysis of a case definition comparing the dose-related effects of vet-
ers' reports of hearing nerve agent alarms and location in the computer-modeled Khamisiyah plume to see if more consistent associations would result. Moreover, studies of gene-environment interactions between the PON1 gene and having heard nerve agent alarms should be performed to reinforce or mitigate our conclusions.

Acknowledgments

The design of and field work for the USMHS was supported by US Army Medical Research and Materiel Command grants DAMD17-97-2-0725 and DAMD17-01-1-0741 and US Department of Veterans Affairs Medical Center, Dallas, Tex. contract VA549-P-0027 (to R.W.H.). 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. No government funding was received for the data analysis or writing of the paper.

A large research team of survey specialists from RTI International contributed importantly to the design and performed the field work for the national computer-assisted telephone interview survey of Gulf War-era veterans in the USMHS. 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. Rick Thompson, Eric Cordell, and Jennifer Escober provided program management at the University of Texas Southwestern Medical Center.

References


Disclosure Statement

Dr. Robert Haley, an internist and epidemiologist employed full-time as a faculty member of the University of Texas Southwestern Medical Center, has directed research on Gulf War illnesses since 1994, published extensively in scientific journals, served on the Department of Veterans’ Affairs Research Advisory Committee on Gulf War Veterans’ Illnesses, and received honoraria and fees for lecturing and consulting on this subject. He directed the USMHS, but his contributions to this paper, including the statistical analyses of the USMHS national database, were performed as part of his regular university faculty activities without external funding support.

Jim Tuite 3rd is an international security and risk management consultant, who was employed as a national security advisor to the chairman of the first US Senate committee to investigate possible causes of Gulf War illness in the early 1990s. From 2002 to 2007, he served on the Expert Advisory Panel to the Department of Veterans’ Affairs Research Advisory Committee on Gulf War Veterans’ Illnesses. Subsequently he was employed as a national security advisor to the President pro tempore of the US Senate and as an expert consultant on Gulf War illnesses and emerging security threats to the US Government Accountability Office and in Gulf War-related litigation. While he compiled much of the intelligence information in the course of his prior official employment, he received no compensation for preparing this paper.

188 Neuroepidemiology 2013;40:178–189
Haley/Tuite


