A Magnetoencephalographic (MEG) Study of Gulf War Illness (GWI)

Brian E. Engdahl, Lisa M. James, Ryan D. Miller, Arthur C. Leuthold, Scott M. Lewis, Adam F. Carpenter, Apostolos P. Georgopoulos

A Magnetoencephalographic (MEG) Study of Gulf War Illness (GWI) has affected many Gulf War veterans. It involves several organs, most notably the brain. Neurological-cognitive-mood-related symptoms frequently dominate and are at the root of chronic ill-health and disability in GWI. Here we investigated the neural mechanisms underlying brain dysfunction in GWI in the absence of mental health disorders.

Methods: Eighty-six veterans completed diagnostic interviews to establish the presence of GWI and assess mental health status. Participants diagnosed with GWI met both Center for Disease Control and Kansas criteria. We studied 46 healthy controls and 40 veterans with GWI without mental illness. They all underwent a resting-state magnetoencephalographic (MEG) scan to assess brain communication based on synchronous neural interactions (SNI; Georgopoulos et al., 2007).

Findings: We found substantial differences in SNI between control and GWI groups centered on the cerebellum and frontal cortex. In addition, using the maxima and minima of SNI per sensor as predictors, we successfully classified 94.2% of the 86 participants (95% sensitivity, 93.5% specificity).

Interpretation: These findings document distinct differences in brain function between control and GWI in the absence of mental health comorbidities, differences that are excellent predictors of GWI.

Funding: U.S. Department of Veterans Affairs and University of Minnesota.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Many Allied military personnel who served in the 1990–1991 Persian Gulf War experience various chronic physical and neurocognitive complaints, now commonly referred to as Gulf War Illness (GWI). At least 25% of Gulf War veterans have been affected by diffuse symptoms such as fatigue, musculoskeletal pain, neurological/cognitive/mood (NCM) complaints, respiratory symptoms, gastrointestinal problems, and rashes (Fukuda et al., 1998; Gray et al., 2002; Kang et al., 2009; Steele, 2000). Several population-based studies have demonstrated that these symptoms occur at significantly higher rates in Gulf War veterans relative to their non-deployed peers and other veteran groups (Fukuda et al., 1998; Kang et al., 2009; Steele, 2000).

However, these symptoms typically do not meet criteria for established medical diagnoses, fueling speculation that they primarily reflect psychological distress and related psychiatric disorders (e.g., Posttraumatic Stress Disorder, PTSD) superimposed on vulnerable organ systems (Engel et al., 2000; Weiner et al., 2011), a supposition that has been refuted (National Research Council, 2010). Brain-based objective indicators of psychiatric and medical disorders are increasingly being identified, allowing a fresh approach to this challenge. Burgeoning evidence supports the utility of magnetoencephalography (MEG) in identifying aberrant and disease-specific neural activity (Engdahl et al., 2010; Georgopoulos et al., 2007, 2010; James et al., 2013; Wilson et al., 2016). One approach focuses on characteristic anomalies in synchronized neural interactions (SNI) derived from task-free MEG. Using that approach, we have demonstrated that SNI anomalies accurately discriminate various brain disorders, including PTSD, multiple sclerosis, Alzheimer’s disease, schizophrenia, Sjögren’s syndrome, temporomandibular pain disorder, and chronic alcoholism from normal healthy brain function, providing compelling evidence of candidate biomarkers (Georgopoulos et al., 2007). Similarly, we hypothesized that GWI could also be characterized by SNI abnormalities, i.e., neural miscommunication patterns. Here, by applying MEG to...
2. Materials and Methods

2.1. Study Participants

A total of 86 Gulf War veterans participated in this study as paid volunteers. VA medical records were reviewed to identify potential participants. Veterans who had completed the Gulf War Registry Examination (Murphy et al., 1999) and did not meet exclusionary criteria were recruited for participation. Exclusionary criteria included cardiac pacemakers or implanted ferrous metal (due to magnetic effects on MEG), central nervous system disorders (e.g., Parkinson’s disease, dementia, cerebrovascular accidents, a history of traumatic brain injury, etc.), lifetime psychotic diagnoses, and current alcohol or drug dependence. Since our participants fulfilled the Kansas GWI case definition (Steele, 2000), they did not suffer from any additional conditions specified as exclusionary by that Kansas GWI criteria. Veterans who might have difficulty with the protocol were also excluded. Study participants provided written informed consent prior to initiating study procedures and participants were compensated for their time. The study protocol was approved by the Institutional Review Board at the Minneapolis VA Health Care System.

All participants completed diagnostic interviews to rule out psychiatric diagnoses and underwent a MEG scan. The Clinician-Administered PTSD Scale for DSM-IV (CAPS; Blake et al., 1995) was administered to evaluate current PTSD diagnostic status. Non-PTSD Axis I diagnostic status was determined using the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID; First et al., 2002). SCID screening questions were administered to all participants; positive screening items were further evaluated with the relevant SCID module. All participants completed a GWI symptom questionnaire developed for use in Kansas Gulf War veterans (Steele, 2000) that evaluates the presence and severity of 6 kinds of symptoms characteristic of GWI – namely, fatigue, pain, neurological/mood/cognitive, gastrointestinal, skin rashes, and respiratory. Items are rated on a scale from 0 to 3 (absent, mild, moderate and severe, respectively). The questionnaire permits determination of case status according to either the Centers for Disease Control and Prevention (CDC) criteria (Fukuda et al., 1998) or the Kansas GWI case definition (Steele, 2000). Only participants meeting both sets of criteria were included in this study. Participants were thus assigned into the following 2 groups: healthy controls (N = 46) and GWI (N = 40).

2.2. Data Acquisition

As described previously (Georgopoulos et al., 2007, 2010), subjects lay supine within the electromagnetically shielded chamber and fixated their eyes on a spot ~65 cm in front of them, for 60s. MEG data were acquired using a 248-channel axial gradiometer system (Magnes 3600WH, 4-D Neuroimaging, San Diego, CA), band-filtered between 0.1 and 400 Hz, and sampled at 1017.25 Hz. Data with artifacts (e.g. from non-removable metal or excessive subject motion) were eliminated from further analysis.

2.3. Data Analysis

2.3.1. General

Standard statistical methods were used to analyze the data, including analysis of covariance (ANCOVA), two-sample Kolmogorov-Smirnov test, binary logistic regression, etc. The following packages were employed: IBM-SPSS statistical package, version 23, Matlab (version R2015b), and ad hoc Fortran computer programs employing the International Mathematics and Statistics Library (IMSL; Rogue Wave Software, Louisville, CO, USA) statistical and mathematical libraries.
3.2. Group Differences in SNI

The SNI distributions differed significantly among the 2 groups \( (P < 0.025, \text{Kolmogorov-Smirnov test}) \). The parametric ANCOVA carried out for each sensor pair yielded results that helped identify differential involvement of neural interactions. The basic finding was that, for a given sensor, the Group F-statistic in the ANCOVA (essentially, a signal-to-noise ratio) varied between that sensor and the remaining 247 sensors (this sensor’s “domain”) in a graded fashion, an example is illustrated in Fig. 1. When the F-values for each sensor were rank-ordered, aligned to their maximum, averaged across sensors, and plotted against their rank, an orderly decline in the average F-value was observed with small variation within each rank (Fig. 2). This indicates a very similar distribution of the F-values from sensor to sensor, although the spatial (in sensor space) distribution per sensor was quite distinct. The relative importance of each sensor in distinguishing between control and GWI groups was visualized by a heatmap (Fig. 3) where the color in a given sensor location was varied according to the maximum F-value of that sensor, among the 247 F-values associated with it. It can be seen that major effects were centered around the cerebellum and frontal lobes bilaterally, and frontal lobe midline.

3.3. Participant Classification

The finding that ANCOVA-based group differences varied within a sensor SNI domain in a graded fashion suggested that the variation of SNIs themselves may be good predictors of group assignment. We investigated this possibility by assessing, first, the distribution of SNIs within a sensor domain. We found that SNIs varied in a graded fashion, as illustrated for a specific sensor in Fig. 4. Two characteristic values of these SNI distributions are their highest (maximum) and lowest (minimum) values, as seen in Figs. 4 and 5. Therefore we used those two values per sensor as predictors (total predictors = 248 sensors × 2 predictors per sensor = 496) in a stepwise binary logistic regression where the group membership of a participant (control, GWI) was the dependent variable. We carried out this analysis using the Forward...
Conditional option in the IBM-SPSS statistical package (version 23). This analysis identified 9 predictors which yielded a 94.2% correct classification (Table 1). The sensitivity was 95% (38/40) and the specificity 93.5% (43/46). Fig. 6 shows the Receiver Operating Characteristic (ROC) curve; the area ± SE under the curve was 0.942 ± 0.029 (P = 1.8 × 10^{-12}). We carried out two validation procedures of this classification. In the first, we selected randomly 25 participants from each group and applied the binary logistic regression analysis using the 9 predictors above on 1000 such random samples. The mean ± SEM correct classification rate was 99.5 ± 0.063%; the median was at 100% correct. In the second analysis, we applied the logistic regression analysis to 1000 random bootstrap samples of 46 control and 40 GWI participants. The mean correct classification rate was 99.60 ± 0.047%; the median was at 100% correct.

Finally, it should be pointed out that the 9 predictors above were minima or maxima of 9 specific sensors. Thus, although those 9 sensors were the same across brains, the specific minimum or maximum value for a specific sensor could refer to the SNI of the pair between that (fixed) sensor-predictor and any of the remaining 247 sensors. Now, we found that sensors involved in those minimum or maximum SNIs (for each one of the 9 sensor-predictors) were very similar from brain to brain. Fig. 7 outlines the approximate SNI territories of those sensor pairs superimposed on the similarly scaled heatmap of Fig. 3. It can be seen that these territories roughly correspond to the areas in Fig. 3 that showed strong Group effects in the ANCOVA.

### Table 1

Participant classification results yielded by the binary logistic regression. (see text for details.)

|                | Predicted | Total |
|----------------|-----------|-------|
|                | GWI       | Control |
| Observed       | 38        | 2     | 40 |
| Control        | 3         | 43    | 46 |
| Total          | 41        | 45    | 86 |

B. Analysis of the two-way classification table

| Test                          | Value | DF | Significance (2-sided) |
|-------------------------------|-------|----|------------------------|
| Pearson Chi-square (corrected for continuity) | 63.6  | 1  | P = 1.49 × 10^{-15}    |
| Fisher’s exact test           |       |    | P = 2.19 × 10^{-18}    |

C. Mantel-Haenszel common odds ratio estimate

| Estimated odds ratio (α) | ln (α) | SE of ln (α) | Asymptotic significance (2-sided) |
|--------------------------|--------|--------------|----------------------------------|
| 272.3                    | 5.61   | 0.94         | P = 2.41 × 10^{-15}              |

4. Discussion

Here, using MEG, we investigated dynamic brain function in control and GWI participants in the absence of mental illness. Specifically, we tested the hypothesis that GWI could be characterized by anomalies in synchronized neural activity, as assessed by the SNI test (Georgopoulos et al., 2007). Our results provide strong evidence for this hypothesis at the group level and, moreover, extend the application of the SNI test to the classification of participants based on the variation (maxima and/or minima) of SNI distributions of individual sensors.

4.1. SNI Anomalies in GWI

We have previously demonstrated the power of SNI derived from task-free MEG in successfully discriminating various brain diseases from healthy brain function. Given that several GWI symptoms involve the brain including memory loss, word finding difficulty, concentration problems, dizziness, tremor, and mood changes, we similarly expected and confirmed SNI differences between Gulf War veterans with and without GWI. Moreover, we demonstrated that the differences were localized primarily around the cerebellum and frontal lobes. Prior studies have identified various brain abnormalities in GWI involving the hippocampus, brainstem, basal ganglia, and frontal lobes (for review, see White et al., 2016); few have investigated cerebellar abnormalities.
and the findings have been mixed. Using magnetic resonance imaging, Rayhan et al. (2013) reported reduced cerebellar volume and functional anomalies in a small sample of GWI veterans. Others have identified enhanced cerebellar activation to noxious heat, presumably reflecting increased pain perception, in some GWI veterans relative to controls (Gopinath et al., 2012). In contrast, researchers using magnetic resonance spectroscopy found no significant differences in metabolites in various brain regions when comparing veterans with GWI to controls (Weiner et al., 2011). The present study adds to this small body of literature and provides evidence of robust resting-state differences in the cerebellum and frontal lobes between veterans with GWI and healthy controls, as identified by MEG.

There is a remarkable parallel between the brain-related symptoms observed in GWI and those observed in the “cerebellar cognitive affective syndrome”: (CCAS; Schmahmann & Sherman, 1998). In the case of CCAS, cerebellar lesions disrupt cerebellar modulation of motor, sensory, cognitive, affective, and autonomic systems, giving rise to executive function impairments, visuospatial difficulties, language deficits, and personality change. Given the similarities between GWI and CCAS at the symptom level, additional research aimed at further investigating cerebellar functional and structural anomalies in GWI and the relation of those anomalies to symptoms characteristic of GWI is warranted.

4.2. Classification

Using properties of the SNIs, specifically the maxima and minima SNI per sensor, we were able to correctly classify 93% of participants as either GWI or control. The excellent results of random permutation and bootstrap tests attested to the robustness of this finding across participants, although a large sample would be needed for more extensive validation. Remarkably, correct classification was obtained with merely 9 (out of 496) possible predictors. We hypothesize that this is due to the kind of predictors used in this analysis, namely that the maxima and minima of interactions (SNIs) of a given sensor with the other sensors, in contrast to previous studies (Georgopoulos et al., 2007, 2010; James et al., 2014) where individual SNIs were used as predictors. Since the current predictors were extracted from SNI distributions (per sensor), it is reasonable to suppose that they are more informative, hence more effective and efficient in classification. It would be interesting to investigate whether such an approach would be successful in other applications, including studies based on fMRI (Christova et al., 2014).

Contributions

BEE, LMJ, and RDM contributed to data collection and clinical evaluation; SML and AFC contributed to neurologic evaluation; APG and JJ contributed to data analysis; APG and LMJ wrote the paper. All authors contributed to extensive editing of the paper.

Declaration of interests

The authors do not report any financial disclosures or conflicts of interest.

Role of the Funding Source

Partial funding for this study was provided by the US Department of Veterans Affairs and the University of Minnesota (American Legion Brain Sciences Chair to APG). The sponsors had no role in the current study design, analysis or interpretation, or in the writing of this paper.

Acknowledgments

This work was partially supported by a service directed grant from the United States Department of Veterans Affairs. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

We wish to especially thank our veteran participants and express our appreciation for the assistance of Kari Resel, Alina Shub and Alexandra Alcorn.

References

Blake, D., Weather, F., Nagy, L.M., et al., 1995. Clinician-Administered PTSD Scale. National Center for PTSD, Boston.
Box, G.E., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.
Christova, P., James, L.M., Engdahl, B.E., Lewis, S.M., Georgopoulos, A.P., 2015. Diagnosis of posttraumatic stress disorder (PTSD) based on correlations of prewhitened fMRI data: outcomes and areas involved. Exp. Brain Res. 233, 2695–2705.
Engdahl, B., Leuthold, A., Tan, H.R., et al., 2010. Post-traumatic stress disorder: a right temporal lobe syndrome? J. Neural Eng. 7, 066005.
Engel, C.C., Li, X., McCarthy, B.D., et al., 2000. Relationship of physical symptoms to posttraumatic stress disorder among veterans seeking care for Gulf war-related health concerns. Psychosom. Med. 62, 739–745.
First, M.B., Spitzer, R.L., Gibbon, M., et al., 2002. Structural Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP). Biometrics Research, New York State Psychiatric Institute, New York.
Fisher, R.A., 1958. Statistical Methods for Research Workers. 13th edition. Oliver and Boyd, Edinburgh.
Fukuda, K., Nisenbaum, R., Stewart, G., et al., 1998. Chronic multisymptom illness affecting air force veterans of the gulf war. JAMA 280, 981–988.
Georgopoulos, A.P., Karageorgiou, E., Leuthold, A., et al., 2007. Synchronous neural interactions assessed by magnetoencephalography: a functional biomarker for brain disorders. J. Neural Eng. 4, 349–355.
Georgopoulos, A.P., Tan, H.R., Lewis, S.M., et al., 2010. The synchronous neural interactions test as a functional neuromarker for posttraumatic stress disorder (PTSD): a robust classification method based on the bootstrap. J. Neural Eng. 7, 016011.
Gopinath, K., Gandhi, P., Goyal, A., et al., 2012. FMRI reveals abnormal central processing of sensory and pain stimuli in ill Gulf war veterans. Neurotoxicology 33, 261–271.
Gray, G.C., Reed, R.J., Kaiser, K.S., Smith, T.C., Gastañaga, V.M., 2002. Self-reported symptoms and medical conditions among 11,868 gulf war-era veterans: the Seabee health study. Am. J. Epidemiol. 155 (11), 1033–1044.
James, L.M., Engdahl, B.E., Lewis, S., 2013. Differences in neural network modulation by trauma between veterans with PTSD and resilient controls. JAMA Psychiatry 70, 410–418.
James, L.M., Belitskaya-Levy, I., Lu, Y., Wang, H., Engdahl, B.E., Leuthold, A.C., Georgopoulos, A.P., 2014. Development and application of a diagnostic algorithm for posttraumatic stress disorder. Psychiatry Res. 213, 1–7.
Kang, H.K., Li, B., Mahan, C.M., et al., 2009. Health of US veterans of 1991 gulf war: a follow-up survey in 10 years. J. Occup. Environ. Med. 51, 401–410.
Mahan, M.Y., Chorn, C.R., Georgopoulos, A.P., 2015. White Noise Test: detecting autocorrelation and nonstationarities in long time series after ARIMA modeling. Proceedings 14th Python in Science Conference (SciPy 2015), Austin, TX.
Murphy, F.M., Kang, H.K., Dalager, N., et al., 1999. The health status of gulf war veterans: lessons learned from the Department of Veterans Affairs Health Registry. Mil. Med. 164, 327–331.
National Research Council, 2010. Gulf War and Health. Health Effects of Serving in the Gulf War, Update 2009 vol. 8. The National Academies Press, Washington, DC.
White, R.F., Steele, L., O’Callaghan, J.P., et al., 2016. Recent research on gulf war illness and other health problems in veterans of the 1991 gulf war: effects of toxicant exposures during deployment. Cortex 74, 449–475.

Wilson, T.W., Heinrichs-Graham, E., Proskovec, A.L., McDermott, T.J., 2016. Neuroimaging with magnetoencephalography: a dynamic view of brain pathophysiology. Transl. Res. http://dx.doi.org/10.1016/j.trsl.2016.01.007 (epub ahead of print).