SUCCESSFUL ANTIBIOTIC TREATMENT OF THE GULF WAR SYNDROME

A PILOT, RANDOMIZED, PLACEBO CONTROLLED, BLINDED TRIAL

SUCCESSFUL TRIAL OF URINE MICROSCOPY FOR CONTROL OF ANTIBIOTIC TREATMENT OF SYSTEMIC COCCAL DISEASE

Edward S. Hyman M.D, FACP, Louisiana Medical Foundation

William Weiss, Chief (ret) Office of Biometry & Field Studies, National Institute-Neurological Disorders & Stroke

Quentin B. Deming M.D., Jacob A. & Jeanne E. Barkey Professor of Medicine, Emeritus, Albert Einstein College of Medicine

* Author responsible for correspondence
Quentin B. Deming, M.D.
Abstract

Background

Gulf War Syndrome has occurred in a major cluster in thousands of veterans. The etiology is unknown. An infectious etiology has not been generally accepted. The observation, by urine microscopy, of increased excretion of Gram positive cocci by patients led one of us (E.S.H.) to an empirical trial of antibiotics. Its apparent success in 22 patients prompted this trial.

Methods

36 sick veterans were randomized into a placebo-controlled trial with independent, blinded monitoring of the clinical outcomes. The selection of antibiotic regimens was guided by suppression of the excessive excretion of Gram positive cocci and their empty cell walls in the urine as well as clinical responses.

Results

Compared to placebo: fatigue, by one test, was reduced 4-fold, p=0.0074, by another test, 10-fold, p=0.0002. Headaches were reduced from 12.5 to 2.5 per month, p=0.0007. Median “Quality of Life” score was increased from 27.5 to 44.5, p=0.0016. Reduction of pain and improvement of sleep showed “non-significant” p values of 0.038 and 0.061. No effect on cognition was demonstrated by the Neuropsychological impairment index. There were no significant adverse effects.

Conclusion

This trial, clearly demonstrated a response to the antibiotics employed which was both clinically and statistically highly significant. No treatment for this disease has previously been proven effective. A bacterial basis for the syndrome, though not proven, is supported. These striking results should encourage validation by further studies which could employ standardized versions of the antibiotic intervention.

The successful trial also serves as a validation of the value of urine microscopy for both diagnosis of systemic coccal disease and control of antibiotic therapy.
BACKGROUND

According to the Institute of Medicine, “As many as 100,000 of the 700,000 troops who served during Operations Desert Storm and Desert Shield in 1990 and 1991 have reported symptoms ranging from chronic fatigue, muscle and joint pain, memory loss, balance disturbances, sleep disorders, depression, chronic diarrhea and concentration problems.”[1] Similar pathology has been reported in the troops from 5 countries. According to the Department of Veterans Affairs over 180,000 Gulf War veterans’ claims for disability have been granted; triple the rate of WW2 or the Korean war.

This major cluster of a syndrome grossly affecting the quality and some times the duration of life has acquired the name Gulf War Syndrome (GWS). A Presidential Commission was established to study it[2]. This commission raised the possibility of exposure to various toxins, experimental vaccines or psychological stress as possible causes. The Institute of Medicine was directed to study the role of toxins[3] It was unable to find convincing evidence of such a relationship. Neither report offered detailed analysis of the possibility that the syndrome might be of bacterial origin. Two observations suggest to us the importance of reconsidering an infectious etiology. First: in a beautifully conducted epidemiologic study,[4] Lea Steele clearly demonstrated both a temporal and geographic focus of highest incidence of the syndrome, which is common with infectious outbreaks, and second, there have been many claims by affected veterans that their family members were developing similar symptoms. Transmission to family members seems to us inconsistent with an etiology based on exposure to external toxins or immunizations.

Background for the hypothesis about etiology and the selection of the treatment.

The concept that chronic diseases of unknown etiology may be manifestations of occult chronic bacterial infections has been confirmed with the elucidation of such illnesses as Lyme arthritis and peptic ulcer. Its relevance to Gram positive coccal infections, goes back at least as far as Billings in 1915 [5], through Coburn’s [6] explication of acute rheumatic fever and was explored with more modern bacteriological methods by Costerton[7] as recently as 1999. Its confirmation in specific instances has seen the removal from the etiology unknown or “functional” category of
such illnesses as rheumatic fever, Lyme arthritis (with its fatigue), and peptic ulcer. One of us (ESH) has developed an approach to diseases of this type over a period of 20 years.

In 1992 he published a method for preparing stained preparations of urine sediment. This differs from the conventional method in two ways. First, it employs centrifugation at 11,000 x G for 10 minutes. This sediments fragments of bacterial cell wall which are missed at the slow centrifuge speeds (~400xG for 5 minutes) used in clinical laboratories. Second, a chemical method is used which bonds the sediment to the glass slide, so that it is not washed off during staining, as most is during the standard procedure. Conventional staining of sediments prepared with this technique reveals in urines (which are reported as normal from clinical laboratories) Gram positive cocci, cell walls of degenerating cocci, as well as tiny, encapsulated Gram positive cocci which may appear either free, within leucocytes or within casts. While this method reveals a rare Gram positive coccus and some coccal cell walls in almost all urines, their numbers are grossly increased in some diseases and the tiny encapsulated cocci are unusual in the absence of disease.

In 1993 he published a basic mathematical synthesis which established that the presence in urine of steady numbers of organisms, of types which grow slowly if at all in urine, reflects their passage through the blood stream to the urine and measures the magnitude of that transit. Zierdt showed such organisms grow in only a minority of cases if blood from healthy humans is cultured aerobically and anaerobically.

In 1994 his paper “A Urinary Marker for Occult Systemic Coccal Disease” described the increased numbers of Gram positive cocci and shells of Gram positive cocci in the urine of patients with a variety of diseases, often classified as auto-immune, for which a streptococcal cause had been sought but not found. He suggested a related pathogenesis for these diseases. One such disease is Chronic Fatigue Syndrome, which is mimicked by Gulf War Syndrome. This ‘marker’ was also found in each case of the Gulf War Syndrome he examined.

He developed a clinical technique of using repeated estimation of the numbers of cocci excreted during the course of treatment as an in vivo guide to the selection and adequacy of dosage of antibiotic therapy. Conventional in vitro sensitivity tests are not applicable. A vast majority of excreted bacteria, even if alive, will not grow in conventional media (many are fragmented, clearly dead). There is often more than one species of bacteria being excreted. An in vitro test of sensitivity
is dependent on culture of the organism (not possible if it is dead), gives no evidence of penetrance of the antibiotic to occult sites, and measures the sensitivity in non physiological milieu. The simple, rapid urine procedure can be performed repeatedly during the course of treatment and reflects the effect of therapy on the bacteria, not in culture but in situ and if more than one species is present on all of them. Using this method to guide the selection and dosage of antibiotics administered, ESH has treated patients with Chronic Fatigue Syndrome, Arthritis, and other conditions, with results that he and the patients judged to be successful. A placebo controlled trial was clearly necessary.

*Background for application of these methods to Gulf War Syndrome.*

Because of the similarity of the complaints of some sufferers from GWS to those of patients with CFS and other diseases responding to treatment, the urine samples of over 50 veterans suffering from GWS were examined. Each showed a high rate of excretion of coccal forms and degenerating coccal forms. By 1994 he had treated 10 afflicted veterans (and 10 wives and 2 children, who after several months delay had insidiously developed similar complaints), with relief of symptoms in all and restoration in some of employability and work status. Each of the 10 veterans had already undergone extensive workups and treatment at an Army or Veterans' facility without relief. Therefore, this randomized, placebo-controlled, blinded trial was designed to test the hypotheses (a) that Gulf War Syndrome is dependent at some stage in its pathogenesis on bacterial infection and (b) that it can be relieved by antibiotic therapy, modified during its administration by the patient’s clinical response and to evidence from the urine of alteration of bacterial load.
Methods

This is not a conventional study of the effect of a chosen drug at a chosen dosage on a defined disease. It is an evaluation of the effectiveness of a method for continuous revision of both drug selection and dosage in controlling a syndrome.

Gulf War Syndrome was treated as a manifestation of what we have called Systemic Coccal Disease.[1] Treatment was based on previous experience with related syndromes and specifically on our experience in the open trial of Gulf War veterans and their families. The selection of antibiotics and dosage at start of treatment was based on that experience. Modification of drug and dose thereafter was based on continuing clinical evaluation of the patient’s response and frequent observation of changes in the excretion of Gram positive cocci (often encapsulated) and their residues in the urine, as a guide to the severity of presumed circulation of what Costerton[7] has called planktonic bacteria. Failure of clinical response or failure to suppress excretion of cocci led to increase in the dosage or modification of choice of antibiotic. Treatment was conducted in New Orleans for a total of 4 months, the first 2-3 weeks of which were in-hospital. The placebo group was treated identically. Matching infusions were given and changes were made at the same intervals. Visits in the hospital or telephone communications with the patients at home (all made by a blinded nurse) were at the same intervals, and oral medication changes were matched. Patients were evaluated at State University of New York at Stony Brook Long Island. (SUNY) for the primary and secondary variables at entry in the study and at its termination. The protocol and informed consent forms were approved by the IRBs of Touro Infirmary in New Orleans and at SUNY. The informed consent form was additionally approved by the Walter Reed Army Medical Center, which retained the role of safety monitoring. Patients were promised that if they were placed in the placebo group they would subsequently be offered the opportunity to receive the treatment. Ultimately they all made that choice. (This work was and is in compliance with the Helsinki Declaration.)
OUTLINE OF STUDY SEQUENCE

Candidates were brought to New Orleans where an initial history, examination and urine examination were performed by ESH. If they appeared to qualify, the study was explained in detail and, if the candidate wished to participate, an informed consent form was signed.

**Phase 1:** baseline evaluation was then conducted by Dr. Lauren Krupp (LK) at SUNY. If the patient was judged eligible, a second informed consent was signed and the patient was then randomized, by Kunitz & Associates in Maryland, to a treatment or placebo group.

**Phase 2:** approximately 3 weeks of intravenous therapy or matching placebo was given in-hospital in New Orleans. The duration of the study for each patient was 17 weeks from the first day of this “treatment”.

**Phase 3:** oral medication or placebo was then provided at home or at work until the 13\(^{th}\) week, when,

**Phase 4:** he/she returned to New Orleans for 5 days of out-patient intravenous treatment or placebo.

**Phase 5:** the patient returned home (still on oral management) until the 17\(^{th}\) week, at which time “treatment” was discontinued for 24 hours and

**Phase 6:** she/he returned to SUNY for final evaluation.
Flow Diagram of Protocol

Recruitment & Informed Consent (New Orleans)

Pre treatment Evaluation and conformity to admission criteria (SUNY Stony Brook)

Randomization (Kunitz & Associates-Maryland)

Hospitalized IV Rx, 2-3 weeks (New Orleans) *blinded*

oral Rx, 2 months (Home or duty) *blinded*

IV Rx, 5 days (New Orleans) *blinded*

Oral Rx, 1 mo, (Home or duty) *blinded*

Final Evaluation (SUNY Stony Brook) *blinded*

*KAI breaks code & sends data to statistician*
Recruitment & Assignment

57 veterans of the Gulf deployment who responded to internet posting of the study were screened and found to be suffering from measurable fatigue, cognitive dysfunction and somatic pain. Each showed excessive excretion of Gram positive cocci, deformed cocci or degenerating coccal forms in his or her urine.\textsuperscript{[11]} 38 of the 57 volunteered to be randomized into a treatment or a placebo group. Of the 38 patients, 36 had previously been extensively evaluated and treated at DOD or Veterans facilities, without relief. The other two had been deployed to the Gulf as civilian contractees and had been evaluated and treated by private physicians, without success.

Inclusion Criteria

The candidate must have been deployed to the Gulf. His/her symptoms must not have pre-existed the deployment and must have occurred by the end of 1993 (within 2 years of deployment). There must have been no other evident explanation for the symptoms. Each patient had to have both primary conditions, Fatigue and Impaired Cognitive Processing, plus the secondary condition, Somatic Pain, and at least one of the other secondary conditions, Impaired Quality of Life, Headache, Sleep Disorder, Diarrhea. The patient’s urine had to show abnormally increased excretion of Gram positive cocci or degenerated coccal forms\textsuperscript{[11]}. (This was found in every candidate studied).

After the patients had agreed to the terms of the study and had been determined to have met the inclusion criteria, their numbers were sent to an independent data management company, Kunitz and Associates in Maryland. Under their direction, a computer randomization procedure allocated the patients to placebo or treatment group in randomized groups of four. Thirty-six comprised the evaluable cohort, and the determination of treatment efficacy was based on the data from those patients. The randomization of the last of the 36 patients was reversed to match the groups at 18 each.

Two additional patients were randomized, but shortly thereafter were considered not to have met the requirements for entry into the study. (one because of a diagnosis of rheumatoid arthritis, the other of drug addiction). These 2 patients, though randomized into the placebo group, received treatment-group medication, and their data, together with those of the other 36 patients, comprised a
secondary intent-to-treat cohort. The latter 2 patients in the intent-to-treat cohort, though they had been treated, were analyzed as cases in the placebo group for purposes of comparison.

Blinding

Blinding for this study presented an unusual problem. The therapeutic method being tested required the treating physician to repeatedly vary drugs and dosages administered, on the basis of changing findings in the urine and/or changing responses of the patient. Therefore, the treating physician and the iv pharmacist could not be blinded. All other personnel at all sites were blinded. Precautions were taken to prevent either harm to the patient or breaking of blinding. The pharmacist assigned only an 8 digit number to each medication or placebo. Dr. Hyman made one brief hospital visit daily for safety monitoring and therapeutic decision making. He made an identical visit to a patient in the other group. He attempted to handle all patients identically and was repeatedly accompanied by a blinded colleague who tried to observe any accidental hints. None were perceived. All other patient contacts were made by a blinded nurse, who saw to the patient’s needs in the hospital, observed, questioned and recorded all changes, including possible adverse events. After discharge she maintained all contacts with the patient, by phone and mail, sent pills (matched in appearance) as needed which were provided to her numbered not labeled, by the pharmacist. The nurse also monitored the patient diaries. The patients were blinded. All evaluations both before and after “treatment” were made by the blinded physicians and staff at the State University of New York (who had no contact with the staff at New Orleans) The entire staff at the Louisiana Medical Foundation, was blinded.. Kunitz and Associates Inc. (in Maryland) provided the data management and processing center and performed the patient randomization.
Treatment

Hospital Phase

A continuous intravenous infusion was maintained for the duration of hospitalization which extended from 14 to 25 days (mean 17.7). The duration was determined by the patient’s clinical response and by clearing of the urine of Gram positive cocci. The duration of hospital stays for placebo patients was adjusted to duplicate the stays of patients in the active treatment group. The usual starting infusion, in a 575cc bag, contained: clindamycin 5.4 Gms, cefazolin 2 Gms, ceftizoxime 1 Gm. Started slowly, the rate of infusion was progressively increased. Both placebo and treatment patients were evaluated daily and urine sediments were examined every 2-3 days. In those instances in which the excretion of cocci or their products by a patient in the treatment group, or if he/she was not improving symptomatically, the infusion rate was increased. If improvement did not then result, addition of other agents was tested, selection being based on the therapists accumulating experience, or the patients' previous experience. Whenever a change was made in the infusion for a patient in the treatment group a comparable change was made in the infusion volume and appearance for a patient in the placebo group. The ratio of 5.4-2-1 for clindamycin to 1st generation cephalosporin to 3rd generation cephalosporin was maintained unless the clindamycin dosage was increased above 6 Gms, in which case the cephalosporin dosage was held steady. For the 18 patients in the treatment group, the aggregate dosage of clindamycin ranged from 73 Gms to 223 Gms, with a mean of 127 Gms. The maximum dose/24 hours ranged from 6.1 to 15, with a mean of 9.7. The average dose per day of clindamycin was 7.2 Gms. In those instances in which the excretion of cocci or their products failed to clear and/or the symptoms failed to improve despite increased dosage of the three primary antibiotics addition of other agents was resorted to, selection being based on the therapists accumulating experience or the patients’ previous experiences. Short courses of additional antibiotics included: vancomycin in 9 patients, gentamycin in 4, clarithromycin in 2. Two patients received 2 days of treatment with G-CSF.
Ambulatory phase

Contact was maintained with the patients by a blinded nurse, by phone and diary. Urine samples were examined at intervals of 2 weeks. Changes in medication, if any, were made on the basis of the data so obtained. New, unlabeled, pills were mailed if necessary. Whenever a change was made in the regimen of a patient on treatment, a comparable change was made in the placebos provided to a patient in the placebo group. The psychological or hint effects were thus identical in the two groups. Every patient received oral clindamycin (or placebo) until 24 hours before the start of the final evaluation. The aggregate dose ranged from 76 to 334 Gms with a mean of 162 Gms.

The average dose per day was 1.5 Gms. Ten patients received one or more short courses of clarithromycin (The aggregate dose ranged from 7.5 to 48 Gms, with a mean of 21 Gms). Nine received one or more short courses of cephalexin (The aggregate dosage ranged from 10 to 78 Gms, with a mean of 29 Gms). Five received short courses of a quinolone (The aggregate dose ranged from 3.5 to 10 Gms, with a mean of 8 Gms). Three received a short course of penicillin (The aggregate dosage ranged from 5.5G to 29.5G, with a mean of 15 Gms.).

One month before the final evaluation, all patients were brought back for 5 days of intravenous medication or matched placebo. The infusion over an 8 hour period each day was adjusted on the same basis as infusions during the hospital phase. It ranged from 7.5 to 10.5 Gms clindamycin (with the usual mixture of cephalosporins).

The ambulatory regimen was then resumed until 24 hours before the final evaluation, when it was terminated.
Evaluation Methods:

All urine evaluations were made by ESH in New Orleans. All clinical evaluations before and after treatment, on which results are based, were made by Dr. Lauren Krupp and her associates at State University of New York at Stony Brook. All of them were blinded. A standard set of self-report measures was administered, selected to represent aspects of the patient’s physical, mental and emotional functioning. A standard battery of neuropsychological tests was selected to represent a range of cognitive functioning. (Specific tests and their references are found in Appendix A[30]),

In order to determine the comparability of the 2 treatment groups at baseline, a number of variables were measured that a priori might be associated with the study measures of efficacy. The baseline measures included 4 demographic variables, 4 related to military service, 5 urine assessment variables, a number of other laboratory measurements, and the primary and secondary variables.

The primary variables, upon which the determination of treatment efficacy was based, were:

1. Fatigue-
   a. Modified Fatigue Impact Scale (Fisk\textsuperscript{(app A, 30)})
   b. Fatigue Assessment Inventory;

2. Neuropsychological Impairment Index (which synthesized 30 individual tests).

The secondary variables, to provide additional supportive data, were:

1. Pain,
   a. Visual Analog Scale (McGill)
   b. Dolorimeter\textsuperscript{(app A,30)};

2. Headache, was measured at baseline by present/absent and by number/month. The final assessment of Headache was based on the number in the last month in the study and whether headache prevented work in the last month.

3. Diarrhea,
   a. Frequency
   b. Severity;

4. Sleep disorder;

5. Quality of life\textsuperscript{(app A,30)}. 
Statistical Methods

1. Continuous, normally distributed variables were tested for treatment group differences by two-tailed t-tests. If the variables differed from a normal distribution, they were tested by the Wilcoxon rank sum test [12].

2. Categorical variables, such as race and sex, were tested by a two-tailed, Fisher’s exact test.

3. One of the primary outcome variables, Fatigue, is based on 2 tests; the Modified Fatigue Impact Scale (Fisk), and the Fatigue Assessment Inventory. A combined statistical analysis of these 2 tests was accomplished by a combined Wilcoxon rank sum test [13].

Since the clinical trial has 2 primary variables, Fatigue and Cognitive Function, the required significance level, based on the Bonferroni adjustment [14], was p 0.025 for at least one of the variables to demonstrate treatment efficacy for that variable, in order that the study Type-1 error would remain at 5%.

The sample size of 36 was calculated (using a log rank test of proportions [15]) to achieve a power (1-Type 2 error) of 80% in this trial in order to detect a change from 20% of the placebo patients demonstrating improvement to 80% of the treated patients demonstrating improvement, with a study Type-1 error of 5%, after a 4 month period of treatment.
Results

*Study population characteristics at baseline*

The placebo and treatment groups were compared for age, race, sex, % with college education, military status (% medical leave), military background (% active duty) and time in the Gulf.

None of these baseline variables showed a probability of $p \leq 0.05$ in the differences between the 2 groups of study patients.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>PLACEBO</th>
<th>TREATMENT</th>
<th>TOTAL N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) mean</td>
<td>42.1</td>
<td>39.9</td>
<td>36</td>
</tr>
<tr>
<td>Race, % white</td>
<td>94.1</td>
<td>72.2</td>
<td>35</td>
</tr>
<tr>
<td>Sex, % male</td>
<td>88.9</td>
<td>83.3</td>
<td>36</td>
</tr>
<tr>
<td>Education, % college</td>
<td>70.6</td>
<td>88.9</td>
<td>35</td>
</tr>
<tr>
<td>Military Status, % medical leave</td>
<td>6.2</td>
<td>11.8</td>
<td>33</td>
</tr>
<tr>
<td>Military Background % active duty</td>
<td>82.3</td>
<td>68.8</td>
<td>33</td>
</tr>
<tr>
<td>Time in Gulf, Median (days)</td>
<td>182</td>
<td>197</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 1  **Study population characteristics at baseline**
Nor did any of the 5 urine assessment variables show a probability of $p \leq 0.05$ in the differences between treatment groups.

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Treatment</th>
<th>Probability*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein, %&lt; 2mg/dl</td>
<td>52.9</td>
<td>33.3</td>
<td>0.32</td>
</tr>
<tr>
<td>Gram + cocci</td>
<td>29.4</td>
<td>27.8</td>
<td>1.00</td>
</tr>
<tr>
<td>Abnormal cocci+</td>
<td>64.7</td>
<td>44.4</td>
<td>0.31</td>
</tr>
<tr>
<td>Exploded cocci</td>
<td>82.4</td>
<td>72.2</td>
<td>0.69</td>
</tr>
<tr>
<td>Gram- Rods</td>
<td>11.8</td>
<td>11.1</td>
<td>1.0</td>
</tr>
</tbody>
</table>

*Fischer’s exact test (2-tail)

Table 2  Urine Variables
The patients in each treatment group were queried about exposure to 12 potential hazards. These were chemical warfare/nerve gas, cigarette smoke, DEET-insect repellant, diesel-fueled tent heaters, Iraqi prisoners of war, oil well fire smoke, petroleum-contaminated drinking water, inoculations, pyridostigmine pills, ethanol excess, recreational drug use, flea collars, and pesticides. No differences between the treatment groups in the percent of patients exposed to each hazard were found with a probability of $p \leq 0.05$.

<table>
<thead>
<tr>
<th>Hazard</th>
<th>% Exposure Placebo</th>
<th>% Exposure Treatment</th>
<th>Prob*</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical warfare/nerve gas</td>
<td>41.7</td>
<td>61.5</td>
<td>0.43</td>
<td>25</td>
</tr>
<tr>
<td>Cigarette Smoke</td>
<td>23.1</td>
<td>29.4</td>
<td>1.00</td>
<td>30</td>
</tr>
<tr>
<td>Deet- Insect repellent</td>
<td>30.8</td>
<td>33.3</td>
<td>1.00</td>
<td>30</td>
</tr>
<tr>
<td>Diesel fueled tent heaters</td>
<td>42.9</td>
<td>52.9</td>
<td>0.72</td>
<td>31</td>
</tr>
<tr>
<td>Iraqi POWs</td>
<td>14.3</td>
<td>50.0</td>
<td>0.06</td>
<td>30</td>
</tr>
<tr>
<td>Oil Well Fire Smoke</td>
<td>64.3</td>
<td>82.4</td>
<td>0.41</td>
<td>31</td>
</tr>
<tr>
<td>Petroleum Contaminated H2O</td>
<td>30.8</td>
<td>31.2</td>
<td>1.00</td>
<td>29</td>
</tr>
<tr>
<td>Inoculations (Pyridostigmine)</td>
<td>58.3</td>
<td>94.1</td>
<td>0.06</td>
<td>29</td>
</tr>
<tr>
<td>Ethanol Excess</td>
<td>0.0</td>
<td>0.0</td>
<td>1.00</td>
<td>30</td>
</tr>
<tr>
<td>Recreational Drug Use</td>
<td>0.0</td>
<td>0.0</td>
<td>1.00</td>
<td>30</td>
</tr>
<tr>
<td>Flea Collars</td>
<td>0.0</td>
<td>5.9</td>
<td>1.00</td>
<td>30</td>
</tr>
<tr>
<td>Pesticides, Uniform</td>
<td>45.4</td>
<td>73.3</td>
<td>0.23</td>
<td>26</td>
</tr>
<tr>
<td>Other Exposures</td>
<td>20.0</td>
<td>29.4</td>
<td>0.69</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 3 Exposure to Hazards
**Outcome variables at baseline**

None of the outcome variables demonstrated differences of $p \leq 0.05$ between the treatment and placebo groups at baseline.

<table>
<thead>
<tr>
<th>OUTCOME VARIABLE</th>
<th>PLACEBO</th>
<th>TREATMENT</th>
<th>TOTAL N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisk, mean score (ms)</td>
<td>15.1</td>
<td>14.9</td>
<td>36</td>
</tr>
<tr>
<td>Fatigue Assessment Index (ms)</td>
<td>5.9</td>
<td>5.9</td>
<td>36</td>
</tr>
<tr>
<td>Neuropsych impairment index, median score</td>
<td>-0.72*</td>
<td>-0.60</td>
<td>35</td>
</tr>
<tr>
<td>Sleep Quality, median score</td>
<td>3.5</td>
<td>3.7</td>
<td>28</td>
</tr>
<tr>
<td>Headache, % patients with</td>
<td>88.9</td>
<td>83.3</td>
<td>36</td>
</tr>
<tr>
<td>Median number/month</td>
<td>13</td>
<td>18.5</td>
<td>36</td>
</tr>
<tr>
<td>Diarrhea, % ≥ 1/day</td>
<td>37.5</td>
<td>25.0</td>
<td>28</td>
</tr>
<tr>
<td>severity score ≥ 3</td>
<td>55.6</td>
<td>33.3</td>
<td>36</td>
</tr>
<tr>
<td>Pain, McGill, median score</td>
<td>6.3</td>
<td>6.0</td>
<td>36</td>
</tr>
<tr>
<td>Dolorimeter, median score</td>
<td>0.5</td>
<td>1.5</td>
<td>34</td>
</tr>
<tr>
<td>Quality of Life, median score</td>
<td>20.0</td>
<td>22.5</td>
<td>36</td>
</tr>
</tbody>
</table>

*one outlier excluded

Table 4   **Outcome Variables at Baseline**
Efficacy Evaluation

Primary Variables

Fatigue  Two fatigue tests were administered, the Modified Fatigue Impact Scale (Fisk), and the Fatigue Assessment Inventory. There were no significant differences at baseline between the treatment and placebo groups for either of the fatigue tests.

For the Fatigue Assessment Inventory the mean score for the placebo group was 5.7 and for the treated group 3.8, p=0.0005, (Figure 1).

Fig 1  Fatigue Assessment Inventory Scores. The figure compares, for each subject, the score at baseline, before treatment (on the X axis), to the final score, after 4 months of treatment, (on the Y axis). The higher the score, the greater the fatigue. A point on the diagonal line indicates no change between the two evaluations. A point to the left of the line indicates the final score was higher than the baseline score (the patient was more fatigued). A point to the right indicates the final score was lower than the baseline score (the patient improved). The lower the point, the freer of fatigue the patient was at the final evaluation.

Though the baseline scores of the two groups were comparable, eleven of the treated subjects showed a more improved score than the best of the placebo patients.
For the Fisk test at the final evaluation, the median score for the placebo group was 13.0 and for the treated group, 8.5, p=0.0047  (Figure 2)

Figure 2: Modified Fatigue Impact Scale (Fisk) Scores

The figure compares, for each subject, the score at baseline, before treatment (on the X axis), to the final score, after 4 months of treatment, (on the Y axis. The higher the score, the greater the fatigue. A point on the diagonal line indicates no change between the two evaluations. A point to the left of the line indicates the final score was higher than the baseline score (the patient was more fatigued). A point to the right of the line indicates the final score was lower than the baseline score (the patient improved). The lower the point, the freer of fatigue the patient was at the final evaluation.

Though the baseline scores of the two groups were comparable, nine of the treated subjects showed a more improved score than the best of the placebo patients.
When designing the trial, we decided that the measures of the primary variables for analysis would be the differences between the final and baseline scores. The reduction in median scores for the Fisk test was 1.5 for the patients in the placebo group and 6.0 for those in the treated group, p=0.0074. For the Fatigue Assessment Inventory, the mean score for the placebo group was reduced by 3.4% and for those in the treated group by 35.6%, p=0.0002. A change of 10% is usually considered clinically important.

In order to calculate a single probability level for fatigue, we used a combined Wilcoxon rank sum test for the fatigue tests, and tested the differences between the final and baseline scores.

Combining the data from both tests, the difference between treatment groups for Fatigue was significant at p=0.0007. (Figure 3)

![Fig. 3 Pre-Post Fatigue Scores - Combined Ranks](image)

**Fig 3** This graph presents data from both the Fatigue Assessment Inventory and the Fisk, in terms of improvement from baseline, the greatest improvement being to the left. The six subjects showing the greatest improvement were all from the treatment group. In the next group of six, four were from the treatment group. In the last group (least improved) five were from the placebo group.
**Neuropsychological Impairment Index**  
The median scores at the final evaluation worsened by 0.17 for the placebo group and by 0.19 for the treated group, \( p=0.96 \). We tested the difference between baseline and final scores for this major variable. The median score for the patients on the placebo arm was 0.25 and for those on the treatment arm 0.28, \( p=0.79 \).

**Secondary Variables**

**Sleep**  
At the final evaluation, we tested the median scores of those patients reporting at least 3 sleep scores. The median sleep score was 3.7 for the placebo group and 4.5 for the treated group, \( p=0.061 \) (Figure 4).

![Figure 4: Sleep Rating Scores](image_url)

**Fig 4**  
Sleep Rating Scores  
The figure compares, for each subject, the score at baseline, before treatment (on the X axis), To the final score, after treatment, (on the Y axis). The higher the score, the better the sleep rating. A point on the diagonal line indicates no change between the two evaluations. 9 of the treatment group showed improvement. 3 of the placebo group did
**Headache**  At the final evaluation, the median number of headaches in the last month was 12.5 for the placebo group and 2.5 for the treated arm, $p=0.0007$ (Figure 5).

![Graph showing the frequency of headaches at baseline and final evaluation.](image)

One placebo patient, outside the square, reported 90 headaches per month at baseline and 92 headaches at final evaluation.

**Figure 5: Headaches per Month**

**Fig 5 Headaches Per Month**

The figure compares, for each subject, the frequency of headaches at baseline, before treatment, (on the X axis) to the frequency of headaches in the final month of treatment, (on the Y axis). The higher the score, the more frequent the headaches. A point on the diagonal line indicates no change between the two evaluations. A point to the left of the line indicates the headaches were more frequent at the final evaluation. A point to the right of the line indicates the frequency of headaches had decreased, (the patient improved). The lower the point, the fewer the headaches.

At the final evaluation 11 of the treatment patients had shown improvement and 7 of these were essentially headache free. 6 of the placebo patients had shown improvement and of these 1 was essentially headache free.

With regard to the question of whether the headache(s) prevented work or normal activity, 58.8% of patients in the placebo group answered “yes”, as compared to 40.0% of those in the treatment group, $p=0.44$.  

23
**Diarrhea**  In patients reporting diarrhea, the frequency of diarrhea was greater than intermittent in 36.4% of the placebo group and in 50.0% of the treated group, p=0.67. Those reporting severity of diarrhea with a severity score ≥3 were 27.8% of the placebo group, and 11.1% of the treated group, p=0.40.

**Pain**  Pain was evaluated by 2 tests, the Visual Analog Scale (McGill), and the Dolorimeter. For the McGill test, the median score was 5.8 for patients in the placebo group, and 5.4 for those in the treated group, p=0.038 (Figure 6).

![Image of a scatter plot showing pain scores at baseline and final evaluation, with placebo and treatment groups differentiated by symbols. The plot illustrates the change in pain scores over time, with a diagonal line indicating no change.](image)

**Figure 6: Pain: Visual Analog Scale (McGill) Scores**

The figure compares for each subject, the score at baseline, before treatment (on the X axis), to the final score, after 4 months of treatment, (on the Y axis). The higher the score, the greater the pain. A point on the diagonal line indicates no change between the two evaluations. A point on the left of the line indicates the final score was higher than the baseline score (the patient was having more pain). A point to the right of the line indicates the final score was lower than the baseline score (the patient was having less pain). The lower the point, the freer from pain the patient was at the final evaluation.

For the Dolorimeter test, 33.3% of the patients in the placebo group scored ≥1, while 29.4% of those in the treated group scored ≥1, p=1.00.
**Quality of Life** (SF-36) This is a test of general health perception, with scores ranging from 0 to 100. Scores of 20-30 are considered extremely low. The median baseline scores here were 20 and 22.5. At the final evaluation, the median score for patients in the placebo group was 27.5 and for those in the treated group, 44.5, p=0.0016 (Figure 7).

*Fig 7: Quality of Life Scores*

The figure compares, for each subject, the score at baseline, before treatment (on the X axis), to the final score, after 4 months of treatment, (on the Y axis). The higher the score, the better the quality of life. A point on the diagonal line indicates no change between the two evaluations. A point to the left of the line indicates the final score was higher than the baseline score (the patient’s quality of life improved.) A point to the right of the line indicates the final score was lower than the baseline score (the patient’s quality of life declined).

Though the baseline scores of the two groups were comparable, ten of the treated subjects showed a better final score than the best of the placebo subjects. An improvement of 10 points or more would be clinically important (App A 300)
Conclusions

We conclude from the above tests: 1) that the symptoms of fatigue, as measured by a combined test (p=0.0007) and separately by the Fisk (p=0.0074) and Fatigue Assessment Inventory (p=0.0002), were demonstrably diminished in patients on the treatment as compared to those in the placebo group and 2) that the treated patients had fewer headaches (p=0.0007) and an improved quality of life (p=0.0016). While we were not able to demonstrate statistical significance in reduced pain symptoms as measured by the McGill Scale (p=0.038) and improved sleep (p=0.061), the patients in the treated group showed sufficient improvement in both of these symptoms to warrant further investigation. Although patients in the treatment group believed they experienced improvement in memory and attention, the Neuropsychological Impairment Index measured at State University of New York showed no effect of treatment (p=0.79).

Comparison of Evaluable and Intent to Treat Cohorts

The evaluable cohort consisted of 36 patients, whose data were used to determine the efficacy of the treatment. Two additional patients were quickly determined not to have met the criteria for inclusion in the trial. (One because of a diagnosis of rheumatoid arthritis, the other of drug addiction). They participated in the trial process, but on treatment rather than on placebo to which they had been randomized. The 36 patients plus the two excluded-but-randomized, patients formed the intent-to-treat cohort of 38 patients. Because both patients who received treatment improved, the differences between placebo and treatment groups in the intent-to-treat analysis were slightly, but not meaningfully, less than in the evaluable cohort.

Adverse Events

There was no serious adverse event in either the treatment or the placebo group. There were 241 minor events (i.e. Temporary complaints of something not apparently related, from backache to abrasions) in the placebo group and 275 in the treatment group. There were no significant differences between the groups for either the whole list or any individual complaint.
Discussion

**Implication of the results for the validity of the method of urine microscopy for diagnosis and for control of therapy.**

Having found increased excretion of cocci in the urines of over 50 veterans complaining of the Gulf War Syndrome and of 12 of their family members who were not deployed, ESH proposed that the Gulf War Syndrome might have a bacterial basis.

Conventional in vitro sensitivity tests are dependent on a knowledge of which organism is causing the disease, on availability of that organism and on its ability to grow freely in test media. Their usefulness is further dependent on knowledge of the site of infection and the ability to get the selected antibiotic to the relevant site in appropriate concentrations. Such tests could not be used in the case of Gulf War Syndrome, since: 1) all conventional urine or blood cultures were negative. 2) although careful anaerobic cultures at a low oxidation potential, maintained for a number of days, usually did reveal Gram positive cocci, their identity with the cocci revealed in large numbers by the Hyman method of sediment staining [8] has not been established, (although if negative by microscopy they were negative by culture). 3) These cultures grow too slowly to allow testing in a useful time.

If it is assumed that the patient's illness (which is transmissible) is in some manner a result of bacterial action and that elimination of the bacteria is a goal of therapy, then it is desirable to have a measure of whether bacteria are present (alive or dead) and to have a way of measuring whether their number is being affected by therapy.

If it is accepted [9, 11] that the presence of unusual numbers of Gram positive cocci and their shells in the urine is a result of their excretion after passage through the blood stream, a reasonable assumption is that the numbers of bacteria and bacterial products in the urine are related to the numbers of bacteria being produced in or absorbed into the body. In the absence of the traditional pretreatment guide by culture and sensitivity for selection of drug and dose it is nevertheless possible, by microscopic examination of the urine, to rapidly determine whether the therapy chosen is achieving the desired result. This amounts not only to an in vivo test of antibiotic sensitivity of the organism, but additionally a test of whether the antibiotic is reaching the relevant site in
effective concentration. Therapy is guided by both the laboratory and clinical response of the individual patient. The dramatic clinical improvement of most of the patients in this study, which was used as one guide to the adequacy of therapy, is consistent with the validity of this approach. The few failures (3-5 depending on the end point used) are equally consistent. These few include those who received, during the ambulatory phase, the largest amounts of supplementary antibiotics, because their poor response was evident in the urine as well as clinically. This parallelism also supports the validity of the monitoring method. We can not explain this small group of failures. Were the bacteria sequestered in a more protected site, such as within biofilm? Were they different bacteria genetically? Would they have responded had the therapy been even more intensive or prolonged? We do not know. It appears that this treatment is effective for over 80% but not necessarily all patients classifiable as having the Gulf War Syndrome.

The monitoring method also revealed that the suppression of bacterial excretion and by inference of circulating planktonic bacteria, was incomplete. The numbers were effectively suppressed at the end of the in-hospital IV treatment phase, but tended to rise again during the ambulatory treatment phase (though they did not reach the pretreatment levels and the clinical effect of treatment was maintained). This may imply that the treatment, as given, did not completely eradicate the infection at its occult nidus. It raises a question which we can not yet answer about how permanent the clinical improvement may be in the absence of continued suppressive therapy.

Implication of the results for the interpretation of Gulf War Syndrome and its management...

The results demonstrate that this antibacterial regimen controls the symptomatic manifestations of Gulf War Syndrome, for which no treatment has previously been found to be effective. There are published examples of therapeutic effects after antibiotic use (zb tetracycline) which were judged pharmacological rather than antibacterial, but a number of factors in this study make such an explanation seem unlikely. Those patients whose excretion of bacteria persisted in spite of treatment received the highest dosages of antibiotic without benefit. They should have shown the greatest pharmacologic effect, if there were one. Clindamycin has been shown to have a leucocyte stimulating effect, but if relevant, the effect would be as a suppressor of bacteria, not directly on symptoms. The idea that Gulf War Syndrome has a bacterial basis is contrary to prevailing opinion.
Both the Presidential Advisory Committee[2] and the Institute of Medicine[3], though neither could confirm any as the cause, have focused attention on the exposure of veterans to low levels of a variety of toxins. The Advisory Committee said, “..it is unlikely that Gulf War Veterans have infections that have evaded the systematic diagnostic efforts mandated by the standard protocol”.

Richardson and colleagues[16] found that although mental health physicians are more likely to think the syndrome has a physical basis and to consider the possibility of infection, internists are more likely to attribute it to stress and mental disorder. Barky and Borus[17] list Gulf War Syndrome among the many diagnoses which they describe (rightly or wrongly) as “Functional Somatic Syndromes”. In the absence of previous elucidation of etiology or demonstration of effective therapy, this interpretation appears to be the basis of the “supportive” programs provided to veterans by the Army and the Veterans Administration and which have been shown in a large study to be without demonstrable benefit. [18].

Since the results of this study seem to challenge this broad consensus, it is important to consider questions of patient selection, the adequacy of blinding and not just the statistical significance of the changes reported after treatment, but also the magnitude (or clinical importance) of those changes..

*Recruitment:* The long delay (>5 years) since the onset of illness, the long hospitalization required for the trial and our insistence on multi-system manifestation affected recruitment. Our volunteers came from the sickest veterans, who were seeking help on the web after failure of all other treatments, including some who had been on doxycycline (for suspected mycoplasma infection) for as long as a year. They were thus “self selected”, but they were then required to meet selection criteria at both the therapeutic and evaluation sites. While the subjects are thus probably not typical, if these likely sickest patients are helped, the findings may apply with greater force in the population of Gulf War veterans more mildly affected. Further, this group was randomized so the comparison between the placebo and treatment groups was fair.
**Blinding:** These patients had been sick for 5+ years and had undergone multiple therapeutic efforts (which should have exhausted any potential placebo effect) without benefit. Indeed, the Army originally recommended an open trial to us. Nevertheless, we included a placebo group for comparison purposes in designing a proper clinical trial. We also delayed evaluation for 4 months in the belief that any placebo effect from hospitalization would be unlikely to persist that long. Examination of the data from our placebo group reveals little evidence of change from baseline. The continuous individualized revision of therapy required by the protocol, as well as safety and ethical considerations, prevented conventional blinding of the therapist. We therefore designed the trial so that the evaluation team was not only blinded but was at a separate site. At the treatment site, the protocol required minimal contact between the therapist and the patients, and this contact was repeatedly monitored by a blinded colleague trying to observe any accidental hints. He observed none. As the patients completed the final evaluation, both the patients and the examining physician were asked whether they had perceived any hint as to the group assignment of the patients. In every instance such a prior hint, that could skew the results, was denied. We made a major effort to create and maintain blinding. We believe we were successful.

**Magnitude of change:** To obtain p values like those reported from a sample of only 36 patients requires that the effect of treatment must be large. The striking improvement in the patients’ assessment of Quality of Life supports the meaningfulness of the other reported changes. The return of idled veterans to employment or duty attests to the clinical meaningfulness of benefit. At conclusion of the final evaluation and its recording both patient and physician denied having learned or received a hint from any source about the group allocation. Nevertheless, the effects of treatment were so obvious that a correct surmise as to whether the patient had been in the treatment or placebo group was made by the physician in 83% of the 36 and by the patient in 78%.
Interpretation

The evidence therefore does support the hypothesis that the Gulf War Syndrome has a bacterial basis and suggests that the relevant bacteria are Gram positive cocci. It does not prove that hypothesis. One could, for instance, reason that an inflammatory state induced by the illness increased absorption of bacteria from the bowel and thus increased their excretion in the urine and that the action of the medications was as anti-inflammatory rather than as antibiotics. The limited varieties of bacteria cultured, the known actions of Clindamycin and the generalized nature of the clinical response all make such an explanation seem unlikely. If a bacterial infection does explain the syndrome, this study does not reveal the site or sites of the indolent infection. It does not identify the species of Gram positive cocci responsible, nor establish that it is a single strain. These problems remain to be explored. Neither does the study explore the possibility that some unidentified factor in the desert might have increased the susceptibility of the subjects to bacterial infection with these possibly diverse strains of Gram positive cocci. Such possible factors might include: a sustained effect of PTSD on the immune system; a suppressive effect on the immune system of exposure to any of the much studied toxins present in the Gulf theater; the effect of infection with some unidentified virus. The cocci could cause the clinical manifestations and be suppressed by the antibiotics, but might recur if the immune suppression persisted. However, a clearly defined defect in the immune system of these veterans has not yet been established and none of these possibilities would explain transmission of the illness to family members. (Our observation that such transmission is common has not been confirmed statistically).

Though the differences, both clinical and statistical, between the two groups were large, this was a pilot study of a small number and obviously should be confirmed in more subjects. Thousand of veterans continue to request effective treatment from the Army and VA which has not yet been provided. A confirming trial could be conducted without the use of the urine method. Every symptomatic Gulf War Syndrome patient we have examined has shown excessive urinary excretion of Gram positive cocci before treatment. Therefore there would be little risk in admitting patients to a therapeutic trial without this examination. Since application of our method resulted in a fairly similar course of treatment for most of our patients, a standard course of treatment derived from our experience could be selected and followed with clinical observation and no bacteriological
observation. It would probably be less effective for the most refractory patients (for whom our method would lead to higher dosage for longer) and more expensive for the most responsive patients (for whom our method would lead to lower doses for a shorter period) but it would be more familiar methodology for most physicians either to use or to interpret. These results have been so impressive and now, after 9 years of ineffective treatment, placebo seems so irrelevant, that such a trial could and probably should be open.

**Conclusions**

A randomized, placebo-controlled, blinded, pilot study has shown that an antibiotic regimen, controlled by monitoring excretion of Gram positive cocci, is effective in ameliorating a syndrome which affects thousands of Gulf War veterans and for which no treatment has previously been proven effective.

The validity and effectiveness of the urine microscopy method ofr diagnosis and for control of treatment has been confirmed.

The hypothesis that Gulf War Syndrome is bacterial in origin, though not proven, is supported.

**Competing Interests:** None

**Acknowledgments**

The medical and evaluation skills of Dr. Lauren Krupp and her assistance in interpretation were essential as was the professional randomizing, handling and monitoring of data by KAI. We are grateful to former Congressman Robert Livingston, Lt. General Ronald R. Blanck and Dr. Jay P. Sanford. The support of each was essential to the accomplishment of this study.

We thank: Melinda Spearman R.N., and Mercedes Hassler, who maintained the blinding and cared for the patients and the laboratory; Linda Baskerville for the preparation of the charts; Schein Pharmaceutical of Danbury CT, for the gift of clindamycin capsules and Capsugel
Americas, for the gift of the same empty capsules that Shein uses to load with clindamycin; Delta Airlines for the provision of reduced air fares during part of the study.

**Financial Support** was through a contract between the Department of Defense and the Louisiana Medical Foundation, DADA-15-97-C-0010.

**REFERENCES**

[1] Press release 9/7/00 Committee on Health Effects Associated with Exposures During the Gulf War and Health, Institute of Medicine of the National Academy of Sciences. Gulf War and Health: Volume 1. Depleted Uranium, Sarin, Pyridostigmine and Vaccines.


Nephron; 1994;68: 314-326


pp 43-44 1983

[15] Freedman L S. Tables of the numbers of patients Required in Clinical Trials Using the Log 
Rank Test., Statistics in Medicine vol 1, 121-129, 1982

[16] R.D. Richardson, C. C. Engle jr, M. McFall, K. McKnight, J. K. Boehnlein, S. C. Hunt: 
Clinician Attributions for Symptoms and Treatment of Gulf War-related Health 
Concerns; Arch Int Med. 2002, May 28, 161:1289-94


19) Fisk JD, Pontefract A, Ritvo PG et al. (1994). The impact of fatigue on patients with 

of Psychosomatic Research, 37:753-762.


and validation of a multiple ability self-report questionnaire. Journal of Clinical and 
Experimental Neuropsychology, 16, 93-104.
Other relevant references


Appendix A

Methods Used at State University of New York

Self-Report Measures: A standard set of self-report measures were administered in a single session. The measures were selected to represent aspects of the patients’ physical, mental, and emotional functioning. All materials and questions were kept in the strictest confidence, and were not used for purposes outside the study. Fatigue was assessed with the MFIS-5 Questionnaire and the FAI Questionnaire (Fisk, Pontefract, Ritvo, et al.1994[19]; Schwartz, Jandorf, & Krupp, 1993[20]). Assessment of pain was obtained using the McGill Pain Questionnaire (Melzack, 1987[21]). Assessment of the previous night’s sleep function was obtained using the SMH Sleep Questionnaire (Ellis, Johns, Lancaster, et al., 1981[22]). Assessment of mood was obtained using the CES-D Depression Scale (Radloff, 1977[23]). Physical and medical symptoms were assessed using the MOS-36 Questionnaire and the MASQ Questionnaire (Seidenberg, Haltiner, Taylor, et al[24], 1994; Stewart, Hays, & Ware, 1988[25]).
**Cognitive Measures:** A standard battery of neuropsychological tests was administered in one session, with the measures selected to represent a range of cognitive functions. General intellect was assessed with the Wechsler Adult Intelligence Scale-Revised (WAIS-R) Vocabulary and Block Design subtests (Wechsler, 1981[26]). Verbal learning/memory was measured with an immediate and delayed story recall (Wechsler Memory Scale- Revised or WMS-R Logical Memory; Story A and B counterbalanced across participants; Wechsler; 1987[27]) and the Selective Reminding Test (SRT; Buschke & Fuld, 1974[28]). Visual learning and memory was measured with the Continuous Visual Memory Test (CVMT; Trahan & Larrabee, 1985[29]). Attention was measured with the WMS-R Digit Span and Visual Span (forward and backward scores) (Wechsler, 1987[27]); the Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977[30]); Stroop Interference Test (Stroop; Golden, 1978[31]). Vigil K Continuous Performance Test (CPT; 1995[32]). Psychomotor speed was measured by the WAIS-R Digit Symbol subtest; Trail Making Test Parts A and B (Army Individual Test Battery, 1944[33]); and Finger Tapping Test (Reitan & Wolfson, 1985[34]). To provide a consistent metric for comparison, raw scores were transformed to z-scores based on age-normative data (Wechsler, 1981;[26] Heaton, Grant, & Matthews, 1991[35]; Rao, Leo, Bernadin & Unverzagt[36], 1991 and Masur, et al., 1989[37], respectively).