Coenzyme Q10 Benefits Symptoms in Gulf War Veterans: Results of a Randomized Double-Blind Study

Beatrice A. Golomb  
bgolomb@ucsd.edu  
Departments of Medicine and of Family and Preventive Medicine, University of California, San Diego, La Jolla, CA 92093, U.S.A.

Matthew Allison  
mallison@ucsd.edu  
Department of Family and Preventive Medicine, University of California, San Diego, La Jolla, CA 92093, U.S.A.

Sabrina Koperski  
skoperski@ucsd.edu  
Department of Medicine, University of California, San Diego, La Jolla, CA 92093, U.S.A.

Hayley J. Koslik  
hkoslik@ucsd.edu  
Department of Medicine, University of California, San Diego, La Jolla, CA 92093, U.S.A.

Sridevi Devaraj  
sxdevara@texaschildrens.org  
Department of Pathology and Laboratory Medicine, University of California, Davis, Davis, CA 95817, U.S.A.

Janis B. Ritchie  
jbritchie@ucsd.edu  
Department of Medicine, University of California, San Diego, La Jolla, CA 92093, U.S.A.

We sought to assess whether coenzyme Q10 (CoQ10) benefits the chronic multisymptom problems that affect one-quarter to one-third of 1990–1 Gulf War veterans, using a randomized, double-blind, placebo-controlled study. Participants were 46 veterans meeting Kansas and Centers for Disease Control criteria for Gulf War illness. Intervention was PharmaNord

S.D. is now at Baylor College of Medicine and Texas Children’s Hospital, Houston, Texas.
CoQ10 and Gulf War Illness

(Denmark) CoQ10 100 mg per day (Q100), 300 mg per day (Q300), or an identical-appearing placebo for 3.5 ± 0.5 months. General self-rated health (GSRH), the primary outcome, differed across randomization arms at baseline, and sex significantly predicted GSRH change, compelling adjustment for baseline GSRH and prompting sex-stratified analysis. GSRH showed no significant benefit in the combined-sex sample. Among males (85% of participants), Q100 significantly benefited GSRH versus placebo and versus Q300, providing emphasis on Q100. Physical function (summary performance score, SPS) improved on Q100 versus placebo. A rise in CoQ10 approached significance as a predictor of improvement in GSRH and significantly predicted SPS improvement. Among 20 symptoms each present in half or more of the enrolled veterans, direction-of-difference on Q100 versus placebo was favorable for all except sleep problems; sign test 19:1, \( p = 0.00004 \) with several symptoms individually significant. Significance for these symptoms despite the small sample underscores large effect sizes, and an apparent relation of key outcomes to CoQ10 change increases prospects for causality. In conclusion, Q100 conferred benefit to physical function and symptoms in veterans with Gulf War illness. Examination in a larger sample is warranted, and findings from this study can inform the conduct of a larger trial.

1 Introduction

Approximately 700,000 U.S. troops, plus tens of thousands from the United Kingdom, Australia, Canada, and elsewhere were deployed to the 1990-1 Persian Gulf War. An estimated one-fourth to one-third developed chronic, multisymptom health problems (Binns et al., 2008; Institute of Medicine, 2010; Steele, 2000; Unwin et al., 1999), termed Gulf War illness (GWI), with the presence of multiple symptoms spanning fatigue/sleep, cognition/mood/neurological, pain/muscle function, gastroenterological, respiratory, and dermatologic domains (Blanchard et al., 2006; Fukuda et al., 1998; Steele, 2000). These symptoms have not resolved with time (Binns et al., 2008; Hotopf et al., 2003; Ozakinci, Hallman, & Kipen, 2006). The ground war lasted only four days, and a minority of those deployed saw combat or experienced combat stress; moreover, excess symptoms remained after controlling for combat exposures, posttraumatic stress disorder (PTSD), and psychiatric illness (Ismail et al., 2002; Lange et al., 1999; Proctor et al., 1998; Wolfe, Proctor, White, & Friedman, 1998). Combat stress, though bearing a dose-response relationship to PTSD, does not independently predict GWI (Binns et al., 2008).

Evidence inculpates environmental exposures, particularly acetylcholinesterase inhibitors (AChEi) (Golomb, 2008), the toxicity of which strongly involves oxidative stress (OS) and mitochondrial dysfunction (Milatovic, Gupta, & Aschner, 2006; Pena-Llopis, Ferrando, & Pena, 2002).
Indeed, many of the exposures that Gulf War veterans (GWV) experienced are known to be oxidative stressors (i.e., to promote injury, via reactive oxygen species or (ROS), to proteins, lipids, DNA, and RNA), and to confer toxicity through this means (Golomb, 2012). These include AChEi, specifically pyridostigmine bromide (a carbamate, used as a nerve agent pre-treatment adjunct), organophosphate nerve agent, and organophosphate and carbamate pesticides) (Dandapani, Zachariah, Kawitha, Jeyaseelan, & Oommen, 2003; Golomb, 2008, 2012; Gupta, Milatovic, & Dettbarn, 2001a, 2001b; Hai, Varga, & Matkovics, 1995; John, Kale, Rathore, & Bhatnagar, 2001; Li, Shou, Borowitz, & Isom, 2001; Milatovic et al., 2006; Pazdernik, Emerson, Cross, Nelson, & Samson, 2001; Pena-Llopis et al., 2002; Pena-Llopis, Ferrando, & Pena, 2003a, 2003b; Poovala, Huang, & Salahudeen, 1999; Poovala, Kanji, Tachikawa, & Salahudeen, 1998); but also, if to a lesser degree, reactogenic vaccinations (Clapp et al., 2004); fuels, solvents, and exhausts (relevant to Gulf exposures of oil fires, tent heaters—petroleum-related products) (Blaurock, Hippeli, Metz, & Elstner, 1992; Piotrowska, Dlugosz, & Pajak, 2002); metals and heavy metals (Monnet-Tschudi, Zurich, Boschat, Corbaz, & Honegger, 2006; Olivieri et al., 2002; Poliandri, Machiavelli, Quinteros, Cabilla, & Duvilanski, 2006; Risso-de Faverney, Orsini, de Sousa, & Rahmani, 2004; Wolf & Baynes, 2006) (relevant to depleted uranium, and aluminum in vaccine adjuvants); and radiation, as well as radioactivity (Karslioglu et al., 2005; Kim et al., 2006; Miura, 2004; Park et al., 2006; Shi, Wang, Wang, Zhang, & Zhang, 2005; Wan et al., 2006) (also relevant to depleted uranium). Epidemiological evidence supports a particularly strong and consistent relation of AChEi to GWI (Cherry et al., 2001; Gray, Reed, Kaiser, Smith, & Gastanaga, 2002; Haley & Kurt, 1997; Kang, Mahan, Lee, Magee, & Murphy, 2000; Kroenke, Koslowe, & Roy, 1998; Nisenbaum, Barrett, Reyes, & Reeves, 2000; Schumm et al., 2001, 2002; Unwin et al., 1999; Wolfe, Proctor, Erickson, & Hu, 2002), including a dose-response relationship for the discretized AChEi exposure pyridostigmine bromide (Commonwealth Department of Veterans’ Affairs, 2003; Schumm et al., 2001, 2002; Wolfe et al., 2002), and AChEi are particularly potent oxidative stressors (Gupta et al., 2001a, 2001b; John et al., 2001; Klaidman, Adams, Cross, Pazdernik, & Samson, 2003; Milatovic et al., 2006; Pena-Llopis et al., 2003a, 2003b; Poovala et al., 1999). Underscoring the importance of OS in the toxicity of AChEi—the exposure class most strongly and consistently linked to GWI (Golomb, 2008)—antioxidants given just before or after organophosphate exposure mitigate AChEi lethality and morbidity in animals (Pena-Llopis et al., 2003a, 2003b), and chronic neuropsychological dysfunction in organophosphate-exposed workers correlates not with AChE inhibition but with markers of OS (Rohlman & Lein, 2012). But other non-AChEi exposures more provisionally linked to GWI share induction of OS, if less strongly. In addition, the number of exposures experienced has also emerged as a risk factor (Kroenke et al., 1998), compatible with multiple contributors to cell injury via OS having cumulative impact.
CoQ10 and Gulf War Illness

Figure 1: GRG model of GWI: Proposes central role for oxidative stress and mitochondrial dysfunction (OSMD) and sequelae: Abbreviated pathways and expected coenzyme Q10 effects. DU = depleted uranium; GWI = Gulf War illness; OP = organophosphate; PB = pyridostigmine bromide. See Table 1 for further details on the model.

(A larger number of exposures may also signify greater likelihood of certain key specific exposures.) The exposure associations fit with symptom profiles and variable onset latency comport with OS-induced mitochondrial dysfunction, leading us to hypothesize a role for OS and mitochondrial dysfunction in GWI (Golomb, 2012) (see Figure 1 and Table 1).

Effective treatments for GWI are lacking. Modest benefits have been reported with exercise and cognitive-behavioral therapy (Donta et al., 2003), but effects are small in magnitude and the treatment labor-cost-intensive. In those with GWI who also have sleep apnea, CPAP treatment for sleep apnea confers benefit to symptoms and function (Amin, Gold, Broderick, & Gold, 2011). CPAP treatment has also been shown to benefit these symptoms in persons with sleep apnea who do not have GWI.

Coenzyme Q10 (CoQ10) is the primary endogenous lipophilic antioxidant (Lenaz et al., 2002; Littarru & Tiano, 2007) and an electron carrier conferring bioenergetic support (Crane, 2001; Linnane et al., 2002). Previously it has been shown to improve symptoms in settings of mitochondrial dysfunction (Chan, Reichmann, Kogel, Beck, & Gold, 1998; Linnane et al., 2002). It has benefited GWI-relevant symptoms such as fatigue (Judy, Stogsdill, & Folkers, 1998; Judy, Stogsdill, & Judy, 2002; Langsjoen & Folkers, 1993; Langsjoen, Langsjoen, Willis, & Folkers, 1997; Langsjoen, Langsjoen, Langsjoen, & Lucas, 2005), including chronic fatigue (Judy et al., 1998, Judy et al., 2002; Langsjoen & Folkers, 1993; Langsjoen et al., 1997, 2005)—a condition that arises at elevated rates in GWV (Gray et al., 2002; Steele, 2000; Unwin et al., 1999). Notably in one study, CoQ10 showed the greatest
Table 1: Support for GRG Model.

a. Pathway Elements Represented in Figure

<table>
<thead>
<tr>
<th>Branch of OSMD Model</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exposures → OS</strong></td>
<td>Gulf War exposures promote OS—with particularly strong promotion for the chemical class particularly strongly linked to GWI (Golomb, 2012). Certain exposures/OS mediators (e.g., organophosphates) widely depress antioxidant systems (Altuntas, Delibas, &amp; Sutcu, 2002) enhancing vulnerability to other oxidative stressors. Key GWI exposures/mediators of OS (including organophosphates) can affect nuclear receptors that govern detoxification mechanisms (Kojima et al., 2013); thus boosting OS impact of subsequent exposures; these appear to be affected in GWI (Broderick et al., 2013).</td>
</tr>
<tr>
<td>OS ↔ Mitochondrial dysfunction</td>
<td>OS impairs mitochondrial function (Genova et al., 2004; Huet, Dupic, Harrois, &amp; Duranteau, 2011; Wei, 1998). OS and glutamate excitotoxicity may affect mitochondrial fission-fusion (Nguyen et al., 2011) (altering mitochondrial “dynamics” and morphology) which in turn lead to up-regulation in NMDA (glutamate) receptors and OS (Nguyen et al., 2011). (SOD2 overexpression protects, implicating oxidative mechanisms (Nguyen et al., 2011.).) Mitochondrial dysfunction increases OS (Genova et al., 2004; Wei, 1998).</td>
</tr>
<tr>
<td>OS → Endothelial dysfunction</td>
<td>OS impairs endothelial function (Ashfaq et al., 2008; Haj-Yehia et al., 2002; Jarasuniene &amp; Simaitis, 2003).</td>
</tr>
<tr>
<td>OS → Autoimmune predisposition</td>
<td>OS fosters autoimmune activation/autoantibody production (Burek &amp; Rose, 2008; J. Chen, Gusdon, Thayer, &amp; Mathews, 2008; Griffiths, 2005; Khan, Banga, Mashal, &amp; Khan, 2011; Konno et al., 2013; Ma, Battelli, &amp; Hubbs, 2006; Peters et al., 2009; Profumo, Buttari, &amp; Rigano, 2011; Ruuls et al., 1995; Scott, Williams, &amp; Bolton, 1997; Wu, MacPhee, &amp; Oliveira, 2004), modifying proteins, lipids, etc. in a fashion that renders them more vulnerable to autoimmune attack. Autoantibodies promote inflammation (McCully, 2009).</td>
</tr>
<tr>
<td>OS ↔ Inflammation activation</td>
<td>OS promotes inflammation (Bulua et al., 2011; H. Chen et al., 2008; Cottone et al., 2006; Halliday, 2005; Ma et al., 2006; Peters et al., 2009; Rose et al., 2012). Inflammation in turn can promote OS (Cerutti &amp; Trump, 1991; Hald &amp; Lotharius, 2005).</td>
</tr>
<tr>
<td>OS → Apoptosis/cell death</td>
<td>OS induces apoptotic and necrotic cell death (Aoki et al., 2001; Bhosle, Pandey, Huilgol, &amp; Mishra, 2002; Huet et al., 2011; Li et al., 2008; Nishio, Yoshida, Nishiyama, Hatanaka, &amp; Yamada, 2000; Porras et al., 2003; Ramkumar et al., 2012; Risso-de Faverney et al., 2004; Sastre, Pallardo, &amp; Vina, 2000; Takahashi, Masuda, Sun, Centonze, &amp; Herman, 2004; Van De Water et al., 2004; Xie et al., 2008).</td>
</tr>
</tbody>
</table>
Table 1: Continued.

<table>
<thead>
<tr>
<th>Branch of OSMD Model</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apoptosis → Coagulation activation</td>
<td>Cell death/apoptosis triggers coagulation activation (Reutelingsperger &amp; van Heerde, 1997).</td>
</tr>
<tr>
<td>Apoptosis → Inflammation activation</td>
<td>Cell death triggers inflammation activation (McCully, 2009; Reutelingsperger &amp; van Heerde, 1997).</td>
</tr>
<tr>
<td>CoQ10 → Apoptosis/cell death</td>
<td>CoQ10 reduces and protects apoptosis/cell death (Brancato et al., 2002; Crane, 2001; Di Giovanni et al., 2001; Fernandez-Ayala et al., 2000; Juan et al., 2008; Kagan, Davis, Lin, &amp; Zakeri, 1999; Menke et al., 2003; Papucci et al., 2003).</td>
</tr>
<tr>
<td>CoQ10 → OS</td>
<td>CoQ10 reduces OS (Linnane et al., 2002; Littarru &amp; Tiano, 2010; Rosenfeldt et al., 2002). CoQ10 has antioxidant functions (Littarru &amp; Tiano, 2007) and reduces oxidative and nitrate stress (Kunitomo, Yamaguchi, Kagota, &amp; Otsubo, 2008). CoQ10 up-regulates other antioxidant systems (Lee, Tseng, Yen, &amp; Lin, 2013) and also recycles coantioxidants like vitamin E and C back to the reduced form (Dhanasekaran &amp; Ren, 2005; James, Smith, &amp; Murphy, 2004; Kagan &amp; Tyurina, 1998).</td>
</tr>
<tr>
<td>CoQ10 → Mitochondrial dysfunction</td>
<td>CoQ10 supports energy in settings of mitochondrial dysfunction (Barbiroli et al., 1997; Linnane et al., 2002; Littarru &amp; Tiano, 2010; Rosenfeldt et al., 2002), increasing efficiency of oxidative phosphorylation “independent” of respiratory enzyme deficit (Barbiroli et al., 1997).</td>
</tr>
</tbody>
</table>

b. Pathway Elements Not Represented (NR) in Figure

<table>
<thead>
<tr>
<th>Branch of Model NR</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS → Natural killer cell dysfunction</td>
<td>OS depresses natural killer cell number and activity (Frank et al., 2001; Hansson, Asea, Ersson, Hermodsson, &amp; Hellstrand, 1996; Hellstrand, Asea, Dahlgren, &amp; Hermodsson, 1994; McNulty et al., 2011; Nakamura &amp; Matsunaga, 1998; Peraldi, Berrou, Metivier, &amp; Toubert, 2013).</td>
</tr>
<tr>
<td>OS → Immune dysfunction</td>
<td>OS alters other cytokines and contributes to immune senescence (Houze, Larsson, Hellstrand, &amp; Gustavsson, 1996; Kaltschmidt, Sparna, &amp; Kaltschmidt, 1999; Peters et al., 2009; Wu et al., 2004)—yet OS triggers autoimmune activation (see Table 1a).</td>
</tr>
<tr>
<td>OS → Autonomic dysfunction</td>
<td>OS triggers autonomic dysfunction including alterations in heart rate variability (Chahine et al., 2007; Chuang et al., 2012; Hoeldtke, Bryner, &amp; VanDyke, 2011; Pavithran, Nandeesha, Sathiyaapriya, Bobby, &amp; Madanmohan, 2008).</td>
</tr>
</tbody>
</table>
OS $\rightarrow$ Blood-brain barrier disruption
OS increases blood-brain barrier penetrability (and endothelial cell membrane leakiness, which increases brain exposure to future OS exposures) (Khan et al., 2011; Plateel, Dehouck, Torpier, Cecchelli, & Teissier, 1995; Schleien, Eberle, Shaffner, Koehler, & Traystman, 1994; Shukla, Shukla, Dakshit, & Srimal, 1993; Smith, Andrus, Zhang, & Hall, 1994; Stanimirovic, Wong, Ball, & Durkin, 1995; Zuccarello & Anderson, 1993); so may cell energy deficit (Al Ahmad, Gassmann, & Ogunshola, 2012; Haorah, Knie, Leibhart, Ghorpade, & Persidsky, 2005; Kumar, Mittal, Khanna, & Basu, 2008; Lehner et al., 2011; Lochhead et al., 2010; Plateel et al., 1995).

OS $\rightarrow$ Adrenal hormone alterations
OS and resulting reduced cell energy can trigger stress hormone activation and altered regulation (cortisol supports glucose as a cell energy substrate) (Aschbacher et al., 2013; Ghiciuc et al., 2013; Joergensen et al., 2011; Kraemer et al., 2005; Schmoller et al., 2009). Brain cell death is amplified when cell energy depletion is coupled with glucocorticoids/stress hormones (Sapolsky & Pulsinelli, 1985).

Mitochondrial dysfunction $\rightarrow$ Adrenal hormone alterations/dysregulation
Mitochondria are related to stress hormone dysregulation (Hsu et al., 2005). Adrenal hormone production takes place in the mitochondria. Mitochondria play a role in basal and stimulated stress hormone production (Aupetit, Toury, & Legrand, 1980; Fulop, Rajki, Katona, Szanda, & Spat, 2013; Greengard et al., 1967; Psychoyos, Tallan, & Greengard, 1966; Sewer & Li, 2013; Spat, Fulop, & Szanda, 2012; Tallan, Psychoyos, & Greengard, 1967; Wiederkehr et al., 2011).

OS $\leftrightarrow$ Glial activation
OS precedes and contributes to glial activation (Lee, Kannagi, Ferrante, Kowall, & Ryu, 2009; Miwa, Kubo, Morita, Nakanishi, & Kondo, 2004) in which microglia change from their resting state with a ramified structure to an “activated” amoeboid configuration in response to environmental triggers. Microglial activation can in turn promote OS (Hald & Lotharius, 2005; Reynolds, Laurie, Mosley, & Gendelman, 2007; Todd & Butterworth, 1999) and can cause and perpetuate inflammation (Hald & Lotharius, 2005) by upregulating Cox2 involved in synthesis of inflammatory prostaglandins. (Inflammatory cytokines TNF-α, IL-1β can be directly neurotoxic (e.g.) or enhance NO production; NO can bind to superoxide to produce toxic peroxynitrite.)
Table 1: Continued.

<table>
<thead>
<tr>
<th>Branch of Model NR</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS → White matter abnormalities</td>
<td>OS can conduce to white matter damage/myelin abnormalities (Abdel-Salam, Khadrawy, Mohammed, &amp; Youness, 2012; Abdel-Salam, Khadrawy, Salem, &amp; Sleem, 2011; Casta, Quackenbush, Houck, &amp; Korson, 1997; di Penta et al., 2013; Miyamoto et al., 2013; Munoz-Cortes et al., 2013).</td>
</tr>
<tr>
<td>↓ Cell energy → Oxidative stress</td>
<td>Settings of impaired cell energy increase OS, while correction of this reduces OS (Alonso-Fernandez et al., 2008; Alzoghaibi &amp; Bahammam, 2011; Carpagnano et al., 2003; de Lima et al., 2010; Murri et al., 2009, 2011).</td>
</tr>
<tr>
<td>↓ Cell energy → Glutamate excitotoxicity → OS</td>
<td>Impaired cell energy may promote glutamate excitotoxicity, leading to more OS (Nicholls, 2009).</td>
</tr>
<tr>
<td>↓ Cell energy ↔ Body fluid loss</td>
<td>Impaired cell energy may promote loss of body fluids (and thus blood volume), further impairing cell energy, for instance, via changes in aldosterone, antidiuretic hormone, and atrial natriuretic peptide (Ichioka et al., 1992; Raff &amp; Levy, 1986; Yue, Wang, Xu, Li, &amp; Wang, 2009; Yue et al., 2008). We hypothesize this is adaptive for cardiac function in settings of cell energy insufficiency (analogous to use of a diuretic in patients with heart failure). This may further adversely affect perfusion (and thus cell energy), particularly for cells fed by small capillaries.</td>
</tr>
<tr>
<td>CoQ10 → Mitochondrial function</td>
<td>CoQ10 protects mitochondrial DNA content, excitotoxicity, axon integrity, protect from glial activation (Lee et al., 2014).</td>
</tr>
<tr>
<td>CoQ10 → ↓ Inflammation</td>
<td>CoQ10 reduces inflammation (Kunitomo et al., 2008; Lee et al., 2013; Schmelzer et al., 2008).</td>
</tr>
<tr>
<td>CoQ10 → Cell energy</td>
<td>CoQ10 supports cell energy in settings of mitochondrial dysfunction (Barbiroli et al., 1997; Littarru &amp; Tiano, 2007).</td>
</tr>
</tbody>
</table>

Note: Oxidative stress (OS) is used as a simplification to signify both oxidative and nitrative stress. The list of elements is incomplete, but gives the flavor of the broader implications of this model. CoQ10 = coenzyme Q10; GWI = Gulf War illness; NO = nitric oxide.

benefit among a suite of treatments tried for fatigue (Bentler, Hartz, & Kuhn, 2005), muscle pain, or weakness (Bendahan et al., 1992; Chen, Huang, & Chu, 1997; Kelly, Vasu, Gelato, McNurlan, & Lawson, 2005; Langsjoen et al., 2005; Lodi et al., 2001; Nishikawa et al., 1989; Rosenfeldt et al., 2005); cognitive symptoms (Goda et al., 1987; Judy et al., 2002; Langsjoen et al., 2005), shortness of breath (Langsjoen et al., 1997, 2005), neuropathy symptoms (Langsjoen et al., 2005), exercise intolerance (Judy et al., 1998; Kamikawa, Kobayashi, Yamashita, Hayashi, & Yamazaki, 1985), headaches (Sandor et al., 2005) and quality of life (Bendahan et al., 1992; Cooke et al., 2008;
Judy et al., 1998; Nylander & Malmsten, 1998; Rosenfeldt et al., 2007; Sandor et al., 2005), generally with doses of on the order of 90 to 300 mg/day. It has also benefited cases of fibromyalgia in which CoQ10 levels were documented to be low (Cordero et al., 2011). (Bioavailability and presumably antioxidant-prooxidant balance vary by brand.)

While higher doses have been advocated by some, many antioxidants convert to prooxidant at doses that are “high” (or high for the level of OS), in part by depletion of coantioxidants (Kontush, Finckh, Karten, Kohlschutter, & Beisiegel, 1996; Young & Lowe, 2001). We therefore conjectured that a high-quality-control, high-bioavailability brand of CoQ10 in the 100 to 300 mg/day range might benefit perceived health, function, and symptoms in GWV affected by GWI.

Neural computation is highly energy demanding, as indeed is the brain as a whole. The brain comprises about 2% of body weight yet is responsible for consumption of about 20% of the oxygen (Erecinska & Silver, 2001; Shulman, Rothman, Behar, & Hyder, 2004) and 50% of the glucose used by the body (Fehm, Kern, & Peters, 2006). For this reason, energy support, direct and indirect, including by CoQ10, in settings of hypothesized energy insufficiency, may specifically or particularly confer benefit to the brain.

2 Methods

A randomized, double-blind, placebo-controlled study of CoQ10 100 mg/day (Q100) versus 300 mg/day (Q300) versus placebo was conducted in GWV with GWI. A crossover study was initially planned, but evaluation was modified to focus on the initial parallel design randomized trial (see appendix B).

Participants were 46 community-dwelling southern California GWV meeting Centers for Disease Control and Prevention (CDC) and Kansas criteria (Steele, 2000) for GWI (see appendix A). The study and all study materials were approved by the University of California, San Diego, Human Research Protections Program, as well as the USAMRMC Office of Research Protection. All participants gave written informed consent to participate in the study.

PharmaNord-Myoquinone-CoQ10 or identical placebo softgels from the manufacturer were taken as 1 softgel three times a day for periods of 3.5 ± 0.5 months, with two softgels taken from a larger and one from a smaller container. Depending on randomization assignment, all three (identical-appearing) softgels were Q100, all three were placebo, or one was Q100 and two were placebo softgels. Half of the participants (N = 23) were randomized to placebo and half (N = 23) to CoQ10, with 11 randomized to Q100 and 12 to Q300. Figure 2 shows numbers randomized, as well as dropouts or withdrawals prior to analysis (with reasons for dropout). Of the
11 (Q100), 12 (Q300), and 23 (placebo) randomized, 11 (Q100), 11 (Q300), and 20 (placebo) provided data for analysis.

2.1 Choice of CoQ10 Product and Composition of Placebo

2.1.1 Choice of CoQ10 Product. PharmaNord’s CoQ10 was selected due to: excellent quality control; absence of common ingredients like titanium dioxide or stearates that might have prooxidant or adverse cellular effects; prior evidence of reduction in malondialdehyde (MDA), a marker of oxidative stress (Singh, Niaz, Sindberg, Moesgaard, & Littarru, 2005); use of a softgel capsule, which has shown better bioavailability than hard tablets (Chopra, Bhagavan, Sinatra, & Goldman, 1998); evidence of high bioavailability of PharmaNord CoQ10 against other CoQ10 products in a bioavailability comparison study (Weis et al., 1994); and presence of an existing IND for the PharmaNord myoquinone product.

2.1.2 Placebo Composition. The placebo, provided by PharmaNord and identical in appearance to the CoQ10 softgels, comprised the nonactive ingredients from the CoQ10 softgel. These were capsule shell (gelatin), humectant (glycerol (E 422), purified water), bulking agent (soybean oil, mixture of hydrogenated oils), antioxidant (d-α-tocopherol), and colorant (iron oxide (E 172)).
2.2 Study Recruitment, Run-In, Randomization, Blinding, Study Medication, and Other Medication Use.

2.2.1 Recruitment. Details of recruitment have been published separately (Erickson, Ritchie, Javors, & Golomb, 2013). Multiple approaches were employed, including media posts, talks to clinician groups that included VA physicians, and Internet posts.

2.2.2 Run-In. Participants were seen at a run-in visit at which eligibility was affirmed and a two-week supply of softgels given (placebo run-in).

2.2.3 Randomization and Blinding. Participants who returned for the baseline visit were randomized to placebo or CoQ10. A computer-generated randomization schedule was used to randomize sequential participants to one of four randomization arms: two groups received placebo (or equivalently, one larger group), one received Q100, and one Q300. The randomization schedule was generated by J. O. Denenberg, who has performed statistical work for many UCSD projects. J.O.D. passed the randomization schedule as a single hard copy (to protect from potential electronic access breach) to the individual designated to serve as study pharmacist, who maintained this in a locked cabinet not accessible to others in the group. Sequential participant ID numbers were assigned to the designated softgel sets by the study pharmacist. Sequentially qualifying participants were assigned sequential study IDs and given the corresponding softgel sets under the supervision of the study clinic manager, who supervised assessment for eligibility and enrolled participants. Among study staff, only the study pharmacist, who did not interact with participants, had access to the randomization schedule (which was not on the system network or accessible by any other party, including any party who saw participants) and prepared the softgels for participants according to their randomization assignment. No participants or investigators, no health care providers for participants and no study personnel with any interaction with participants or involvement in participant scheduling or visits, data collection, outcome assessment, laboratory testing, or data entry had any knowledge of or access to randomization assignment.

2.2.4 CoQ10 versus Placebo Allocation and Use. Participants received two containers with softgels: one larger container (with twice the number of softgels) and one smaller container. Depending on the randomization arm, both containers had placebo or both had CoQ10 (providing CoQ10 100 mg tid, totaling 300 mg/day or Q300), or the smaller contained CoQ10 and the larger had placebo (providing CoQ10 at 100 mg/day or Q100). Participants were instructed to take two softgels each day from the larger and one softgel each day from the smaller container three times a day, recommended not close to bedtime. Participants were advised that if they missed one dose,
they could double-up with the next one. CoQ10 can be taken once daily, but divided dosing produces superior blood levels (Singh et al., 2005). Participants were advised to avoid near-bedtime use if possible because activation insomnia can occur with CoQ10, especially at doses of 100 mg or more near bedtime. Our protocol specified that any participant who called with sleep problems would be told to take the CoQ10 earlier in the day.

2.2.5 Concurrent Medications. Participants’ current medications were recorded. Recent (within 3 months) supplementation with CoQ10 precluded enrollment. Participants were asked not to undertake discretionary change in their supplement or medication regimen for the duration of the study, though it was understood that some variance would be unavoidable.

All visits took place at the Clinical Translational Research Institute (CTRI) at the University of California, San Diego. Visits took place at screening/run-in, baseline, and the conclusion of the treatment periods (3.5 ± 0.5 months). Visits were scheduled between 8:30 a.m. and 1:00 p.m., based on participant convenience and CTRI availability.

General self-rated health (GSRH), assessed by written questionnaire, was the specified primary measure: participants self-rated their health as poor, fair, good, very good, or excellent (scored ordinally from 1 to 5). Physical function assessed by the lower-extremity Summary Performance Score (SPS), summing results for timed chair rises, walking velocity, and standing balance three ways, was a secondary objective measure (Guralnik et al., 1994). In light of the small sample, and because participants were on average near the maximum score (10 points on average out of a maximum 12), categories were collapsed to “improvement” versus “not” (binary assessment) for analysis. Individual symptoms were assessed by questionnaire, rated 0 to 10 at baseline, and on follow-up on a 5-point scale from much worse to much better than at the prior visit. CoQ10 levels/CoQ10 effect modification and dose comparison were secondary measures examined. Digit span backward was a specified assessment, to garner objective data on cognition, although the proposal stipulated that self-ratings of cognition were expected to be more sensitive. Rated measures have a wide range of normal and have large and variable training effects. With self-rated cognitive problems, participants can implicitly average recent experience and norm it to prior function. Self-rated cognitive problems have been more sensitive to underlying processes. For example, even in the presence of normal neuropsychological function testing, self-rated memory problems are associated with entorhinal cortex volume loss (Jessen et al., 2006) and predict increased dementia on follow-up (Geerlings, Jonker, Bouter, Ader, & Schmand, 1999; Jorm, Christensen, Korten, Jacomb, & Henderson, 2001; Reisberg, Shulman, Torossian, Leng, & Zhu, 2010; St. John & Montgomery, 2002; Wang et al., 2004). In addition to the specific symptom queries that related to pain (e.g., muscle pain, joint pain, headache),
visual analog scales of pain intensity and pain unpleasantness were assessed.

All eligible and randomized participants with data were analyzed, irrespective of compliance (intent-to-treat).

Effect modification by baseline GSRH (self-rated) was present for symptom change ratings (also self-rated). This was important given baseline GSRH disparities, necessitating adjustment for the interaction and its components for the first-phase symptom-change analysis. The sample size poorly supported these adjustments, leading to unstable estimates and limiting confidence in results.

The profile of GWI has been found to differ, in terms of symptoms and objective markers, for men versus women (Klimas, 2014; Stein et al., 2004), emphasizing the need to assess treatments separately by gender. Sex-stratified analysis was undertaken (hypothesis generating).

Unpaired t-tests were employed for continuous variables. A sign test assessed whether the fraction of symptoms with a favorable direction on treatment versus placebo differed significantly from the null hypothesis. Regression (linear, logistic, or ordinal logit), allowing adjustment for baseline values, was used for outcomes with evidence of baseline disparities across randomization arms or effect interactions based on those disparities. Interpretation of odds ratios with ordinal logit is as follows. A unit change in the predictor variable signifies that the odds for the outcome being in a group that is greater than \( k \) versus less than or equal to \( k \) is the proportional odds times larger (Institute for Digital Research and Education UCLA, 2014). Robust (heteroskedasticity independent or “White”) standard errors (SEs) were used for regression analyses (White, 1980). Regression was also employed to assess for mediation by CoQ10 change. This analysis adjusted for baseline CoQ10 level and baseline values of the outcome variable and excluded participants (from all randomization arms) who were at the ceiling for the outcome at baseline, as the intent was to understand how Q10 change related to outcome change in those who could change. (Floor effects were not observed for these outcomes.) Chi-squared assessed presence/absence of SPS improvement. P-values below 0.05 designated statistical significance. An important purpose of a first treatment study of this kind, in a condition affecting symptoms, function, and well-being, is to ascertain which outcomes appear most promising, and the inclusion of plural outcomes should be viewed in this context. For this reason and since type II error is the greater risk in assessing a promising new hypothesis, it was prespecified that adjustment for multiple assessments would not be performed. Stata 8.0 and 11.0 were used.

3 Results

Participant characteristics are shown in Table 2. Almost all participants characterized their health prior to Gulf War participation as excellent or
Table 2: Baseline Comparability Across Randomization Groups.  

a. Demographics, Nonsymptom Outcomes

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Randomization Arm</th>
<th>P-Value for Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo (N = 23)</td>
<td>Q100 (N = 11)</td>
</tr>
<tr>
<td>Age (years): Mean ± SD (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>48 ± 6.1 (45–50)</td>
<td>50 ± 7.6 (45–55)</td>
</tr>
<tr>
<td>Sex (% male)</td>
<td>91</td>
<td>73</td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>55</td>
<td>73</td>
</tr>
<tr>
<td>Latino</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>African American</td>
<td>9.1</td>
<td>9.1</td>
</tr>
<tr>
<td>Asian</td>
<td>4.6</td>
<td>0</td>
</tr>
<tr>
<td>Native American</td>
<td>4.6</td>
<td>9.1</td>
</tr>
<tr>
<td>Other</td>
<td>9.1</td>
<td>9.1</td>
</tr>
<tr>
<td>Marital (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>8.7</td>
<td>18</td>
</tr>
<tr>
<td>Currently in a relationship</td>
<td>13</td>
<td>9.1</td>
</tr>
<tr>
<td>Currently married</td>
<td>70</td>
<td>64</td>
</tr>
<tr>
<td>Separated</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Divorced</td>
<td>8.7</td>
<td>9.1</td>
</tr>
<tr>
<td>Widowed</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Branch (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Army</td>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td>Marines</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td>Navy</td>
<td>43</td>
<td>36</td>
</tr>
<tr>
<td>Air Force</td>
<td>8.7</td>
<td>0</td>
</tr>
</tbody>
</table>

P-Values for comparison:
- Placebo versus Q100
- Placebo versus Q300
- Q100 versus Q300
Table 2: Continued.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Randomization Arm</th>
<th>P-Value for Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo (N = 23)</td>
<td>Q100 (N = 11)</td>
</tr>
<tr>
<td>Military status in 1990 (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular</td>
<td>87</td>
<td>82</td>
</tr>
<tr>
<td>Reserves</td>
<td>8.7</td>
<td>18</td>
</tr>
<tr>
<td>National Guard</td>
<td>4.4</td>
<td>0</td>
</tr>
<tr>
<td>Officer % (versus % enlisted)</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td>General self-rated health</td>
<td>2.48 ± 0.63</td>
<td>2.00 ± 0.67</td>
</tr>
<tr>
<td>(2.21, 2.75)</td>
<td>(1.55, 2.45)</td>
<td>(2.06, 3.28)</td>
</tr>
<tr>
<td>Summary Performance Score (run-in)</td>
<td>10 ± 1.4</td>
<td>9.5 ± 1.7</td>
</tr>
<tr>
<td>(9.58, 10.8)</td>
<td>(8.32, 10.6)</td>
<td>(8.95, 11.7)</td>
</tr>
<tr>
<td>Coenzyme Q10 (μM/ML)</td>
<td>1.4 ± 0.45</td>
<td>1.6 ± 0.45</td>
</tr>
<tr>
<td>(1.20, 1.59)</td>
<td>(1.29, 1.90)</td>
<td>(1.17, 2.00)</td>
</tr>
<tr>
<td>Backward digit span</td>
<td>6.30 ± 2.32</td>
<td>6.73 ± 1.79</td>
</tr>
<tr>
<td>(5.30, 7.31)</td>
<td>(5.52, 7.39)</td>
<td>(4.59, 7.41)</td>
</tr>
</tbody>
</table>
Table 2: *Continued.*

b. Comparability of Symptoms Across Randomization Groups: Percent Reporting Symptoms

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% Citing Baseline Mean</th>
<th>Randomization Arm</th>
<th>P-Value for Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Placebo (N = 23)</td>
<td>Q100 (N = 11)</td>
</tr>
<tr>
<td>Aches and pains</td>
<td>93</td>
<td>5.7</td>
<td>87</td>
</tr>
<tr>
<td>Joint pain</td>
<td>89</td>
<td>5.4</td>
<td>91</td>
</tr>
<tr>
<td>Tired</td>
<td>76</td>
<td>4.9</td>
<td>70</td>
</tr>
<tr>
<td>Sleep problems</td>
<td>74</td>
<td>6.2</td>
<td>74</td>
</tr>
<tr>
<td>Having to go back and recheck things</td>
<td>67</td>
<td>4.1</td>
<td>64</td>
</tr>
<tr>
<td>Muscle pain</td>
<td>65</td>
<td>4.6</td>
<td>74</td>
</tr>
<tr>
<td>Trouble recalling words or names</td>
<td>65</td>
<td>4.2</td>
<td>61</td>
</tr>
<tr>
<td>Irritable</td>
<td>65</td>
<td>4.9</td>
<td>61</td>
</tr>
<tr>
<td>Impatient</td>
<td>63</td>
<td>5.3</td>
<td>61</td>
</tr>
<tr>
<td>Attention problems</td>
<td>63</td>
<td>4.3</td>
<td>57</td>
</tr>
<tr>
<td>Too little energy to get up and do things</td>
<td>62</td>
<td>4.8</td>
<td>55</td>
</tr>
<tr>
<td>Headache</td>
<td>59</td>
<td>5.0</td>
<td>61</td>
</tr>
</tbody>
</table>
Table 2: Continued.

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% Citing Symptom Overall</th>
<th>Baseline Score(^{a}) (Mean)</th>
<th>Randomization Arm</th>
<th>P-Value(^{b}) for Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Placebo ((N = 23))</td>
<td>Q100 ((N = 11))</td>
</tr>
<tr>
<td>Anxiety</td>
<td>59</td>
<td>5.4</td>
<td>52</td>
<td>64</td>
</tr>
<tr>
<td>Muscle fatigue</td>
<td>57</td>
<td>4.7</td>
<td>65</td>
<td>45</td>
</tr>
<tr>
<td>Fatigue with exertion</td>
<td>54</td>
<td>4.6</td>
<td>48</td>
<td>55</td>
</tr>
<tr>
<td>Ringing ears (tinnitus)</td>
<td>54</td>
<td>5.3</td>
<td>57</td>
<td>64</td>
</tr>
<tr>
<td>Forget where going/what doing</td>
<td>54</td>
<td>3.8</td>
<td>48</td>
<td>55</td>
</tr>
<tr>
<td>Dry skin</td>
<td>52</td>
<td>4.8</td>
<td>48</td>
<td>73</td>
</tr>
<tr>
<td>Cold hands or feet</td>
<td>52</td>
<td>3.9</td>
<td>48</td>
<td>55</td>
</tr>
<tr>
<td>Reading comprehension problems</td>
<td>50</td>
<td>3.5</td>
<td>43</td>
<td>36</td>
</tr>
</tbody>
</table>

Notes: CI = 95% confidence interval; N = number of participants in analysis; Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg. T-tests on the baseline ratings (0–10) of symptoms for those who reported them were also nonsignificant. There was no instance of \(p < 0.05\), for any of the 40 comparisons of either coenzyme Q10 group versus placebo (20 symptoms, assessed against placebo for 2 coenzyme Q10 groups). There was one significant \((p < 0.05)\) difference for a symptom for Q100 versus Q300—for impatience. Borderline significance \((p > 0.05 \text{ but } p < 0.1)\) was present for two additional symptoms for Q300 versus Q100 and for one symptom for Q300 versus placebo. For the featured Q100 versus placebo comparison, for no symptom was there a significant or borderline significant baseline difference.

\(^{a}\)Rating 0–10 for those citing presence of a symptom.

\(^{b}\)Chi square (\(\chi^2\)) for fraction with the corresponding symptom.
Table 3: General Self-Rated Health: Change in GSRH from Baseline to On-Treatment.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Analysis</th>
<th>Coefficient or OR</th>
<th>SE</th>
<th>CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSRH combined</td>
<td>Regression (coefficient)</td>
<td>0.30</td>
<td>0.40</td>
<td>-0.52, 1.12</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>Ordinal logit (OR)</td>
<td>1.88</td>
<td>1.89</td>
<td>0.26, 13.4</td>
<td>0.53</td>
</tr>
<tr>
<td>GSRH men</td>
<td>Regression (coefficient)</td>
<td>0.65</td>
<td>0.30</td>
<td>0.03, 1.27</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Ordinal logit (OR)</td>
<td>6.26</td>
<td>5.57</td>
<td>1.09, 35.8</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Notes: CI = 95% confidence interval; GSRH = general self-rated health; OR = odds ratio; Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg.

Findings: Q300 versus placebo: Not significant: All, \( \beta \) (SE)\{CI\}: +0.15 (0.32) {−0.51, 0.82} \( p = 0.64 \). OR 1.27 (SE 0.98) {0.28, 5.73} \( p = 0.76 \). Men, \( \beta \) (SE)\{CI\}: −0.15 (0.29) {−0.74, 0.45} \( p = 0.61 \). OR 0.67 (0.52) {0.15, 3.07} \( p = 0.60 \). Q100 versus Q300: All, \( \beta \) (SE)\{CI\}: +0.35 (0.48) {−0.67, 1.36} \( p = 0.48 \); OR 2.08 (2.29) {0.24, 18.0} \( p = 0.51 \). Men, \( \beta \) (SE)\{CI\}: +0.91 (SE 0.38) {0.089, 1.73} \( p = 0.032 \); OR 13.9 (SE 14.6) {1.79, 108} \( p = 0.012 \).

\textsuperscript{a}Pretreatment values for placebo versus Q100 differ overall, \( p = 0.0509 \), compelling analysis adjusted for pretreatment values. (Values at baseline also differ for men, \( p = 0.0590 \).)

\textsuperscript{b}Odds ratio interpretation: a unit change in the predictor variable (here, being on Q100 versus placebo) signifies that the odds for GSRH being in a group that is greater than \( k \) versus less than or equal to \( k \) is the proportional odds times larger\textsuperscript{99}. Values at baseline and on treatment, confined to those with on-treatment values (analyzed group). Placebo \(( N = 19 )\): baseline 2.42 (0.63), treated 2.58 (0.84), difference in mean 0.16. Q100 \(( N = 10 )\): baseline 2.00 (0.71), treated 2.50 (1.35), difference in mean 0.50. Q300 \(( N = 10 )\): baseline 2.85 (0.94), treated 3.20 (1.23), difference in mean 0.35.

\textsuperscript{c}Male sex was a strongly significant predictor of GSRH change \( ( p = 0.008 ) \), prompting analysis by sex.

very good (96%), while almost none reported excellent or very good health now (6%) (see Table 2a). Baseline comparability across treatment arms was supported for most variables (see Table 2b). However, GSRH differed with borderline significance at baseline for Q100 versus placebo: placebo 2.48 (0.63), Q100 2.00 (0.67), difference 0.48 (SE 0.24) (95% CI 0.00 to 0.96), \( p = 0.05 \).

Because of baseline GSRH disparities, GSRH change was analyzed with regression (ordinal logit) adjusted for baseline values of GSRH. (Results of linear regression are also provided, because many readers are more familiar with this analysis modality, significance is closely similar, and this underscores that findings are robust to the specifics of the analysis.) Run-in and baseline values were averaged for the baseline adjustment, in this and other analyses to minimize regression dilution bias (Emberson et al., 2004; Qizilbash, Duffy, & Rohan, 1991). The change (improvement) in GSRH on Q100 was not significant in the total sample (see Table 3). However, sex differences in drug effects are common (Golomb, 2013; Golomb & Evans, 2008;
including in drug classes relevant to GWV (Menich, Costello, Hoffman, Hershey, & Engler, 2001; Pittman, 2002; Schieszer, 2003) and extending to drug transporters and metabolizing enzymes (Classen & Netter, 1985; Finnen & Hassall, 1984; Ofotokun, 2005; Skett & Paterson, 1985; Xiang et al., 2011). The study sample was predominantly male (85%, comparable to the 83% overall GWV with fatiguing illness who were male (Kang, Natelson, Mahan, Lee, & Murphy, 2003)), and male sex significantly predicted GSRH change in the total sample ($p = 0.031$). Sex-stratified analysis de facto focused on men because there were too few women for separate female-specific analysis. In men, 25 of the 29 participants in the Q100 versus placebo comparison, Q100 significantly predicted benefit to GSRH versus placebo or versus Q300 (see Table 3).

Physical function, objectively assessed by SPS, improved from pretreatment ratings in 82% on Q100 compared to 40% on placebo, a 105% relative increase from a 42% absolute increase (chi-squared $p = 0.025$). Women contributed to this benefit. SPS improvement on Q300 was not significant, though the effect on Q100 did not differ significantly from that on Q300, $p = 0.17$.

CoQ10 treatment was associated with significantly increased CoQ10 blood concentrations ($\mu$m/mL) relative to placebo, a finding affirmed by both the change from baseline and by absolute CoQ10 level on treatment (see Table 4).

Figure 3 shows the change in CoQ10 levels in participants in each treatment arm. Average differences in the CoQ10 level change with treatment were qualitatively as expected by dose: no CoQ10 levels declined except among participants randomized to placebo. The largest CoQ10 increases were seen on the Q300 arm. However, on placebo, at least one participant showed a sizeable increase in CoQ10 level, not offset by participants with corresponding magnitude reductions in CoQ10 levels, and the largest CoQ10 increase on placebo exceeded the largest CoQ10 increase on Q100. Some participants in the Q100 and a set of participants in the Q300 arms failed to show material increases in CoQ10 levels (though none showed a decline, as occurred on placebo).

Table 5 shows the relation of CoQ10 change to outcome change for GSRH and SPS, assessed in the total sample and for the Q100-versus-placebo and Q300-versus-placebo comparisons separately. Analyses adjusted for baseline CoQ10 and baseline values of the outcome variable and excluded participants at the ceiling on the outcome at baseline. CoQ10 change was a near-significant predictor of GSRH change in the total sample and separately in analysis excluding the Q100 group. In contrast, significance was markedly attenuated if the Q300 group instead was excluded. CoQ10 change significantly predicted SPS change in the total sample and separately in analysis excluding the Q100 group. In contrast, significance was lost if the Q300 group instead was excluded. Thus, the Q300 group was
Table 4: Secondary Results.

a: Q100 Secondary Results

<table>
<thead>
<tr>
<th>Summary performance score (0–12)</th>
<th>Q100 (N = 11)</th>
<th>Placebo (N = 20)</th>
<th>Difference: Q100 versus Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefit: No</td>
<td>9: 2</td>
<td>Benefit: No</td>
<td>8: 12</td>
</tr>
<tr>
<td>% benefit</td>
<td>82%</td>
<td>% benefit</td>
<td>40%</td>
</tr>
<tr>
<td>Sign of Difference</td>
<td>&gt;</td>
<td>Difference</td>
<td>42%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Chi-squared)</td>
<td>P</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CoQ10 levels (μm/ml)</th>
<th>Q100 (N = 10)</th>
<th>Placebo (N = 19)</th>
<th>Difference, Q100 versus Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change from pretreatment</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean Diff (SE)</td>
</tr>
<tr>
<td>Absolute level</td>
<td>+1.19 (0.69)</td>
<td>0.69, 1.68</td>
<td>+0.24 (0.68)</td>
</tr>
<tr>
<td></td>
<td>2.81 (0.65)</td>
<td>2.34, 3.27</td>
<td>+0.95 (0.27)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Backward digit span</th>
<th>Q100 (N = 11)</th>
<th>Placebo (N = 20)</th>
<th>Difference, Q100 versus Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change from pretreatment</td>
<td>+0.50 (1.58)</td>
<td>-0.56, 1.56</td>
<td>-0.13 (0.71)</td>
</tr>
</tbody>
</table>


Table 4: Continued.

b: Q300 Secondary Results

<table>
<thead>
<tr>
<th>Summary performance score (0–12)</th>
<th>Q300 (N = 11)</th>
<th>Placebo (N = 20)</th>
<th>Difference: Q300 versus Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benefit: No</td>
<td>% benefit</td>
<td>Sign of Difference</td>
</tr>
<tr>
<td></td>
<td>6: 5</td>
<td>55</td>
<td>&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CoQ10 Levels (μm/ml)</th>
<th>Mean (SD)</th>
<th>95% CI</th>
<th>Mean (SD)</th>
<th>95% CI</th>
<th>Mean Diff (SE)</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change from pretreatment</td>
<td>+2.30 (1.97)</td>
<td>0.98, 3.63</td>
<td>&gt;</td>
<td>+0.24 (0.68)</td>
<td>−0.089, 0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute level</td>
<td>3.85 (1.89)</td>
<td>2.57, 5.12</td>
<td>&gt;</td>
<td>1.65 (0.70)</td>
<td>1.31, 1.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Backward digit span</th>
<th>Q300 (N = 11)</th>
<th>Placebo (N = 20)</th>
<th>Difference: Q300 versus Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change from Pretreatment</td>
<td>+0.41 (2.67)</td>
<td>−1.39, 2.20</td>
<td>&lt;</td>
</tr>
</tbody>
</table>

Notes: CI = 95% confidence interval; N = number of participants in analysis; Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg. Summary performance score (SPS) is a 12-point rating measure of physical function. Average ratings were near the ceiling for this instrument, designed with elderly participants in mind (mean 10.31 on a scale of 1–12). In light of proximity to the ceiling, change ratings were binarized for analysis, characterizing performance as improved versus not improved. (Of note, 20% of participants placed on placebo worsened on the SPS versus 0% on Q100.) Q300 difference from Q100: SPS: Chi² = 0.17. CoQ10 level: absolute value: +1.04 (SE = 0.63) (95%CI = −0.28, 2.36) P = 0.12. Change from baseline: +1.12 (SE = 0.66) (95%CI = −0.26, 2.49) P = 0.105. Backward Digit Span: −0.091 (SE = 0.94) (95%CI = −2.04, 1.86) P = 0.92.
Figure 3: CoQ10 change values on placebo, Q100, and Q300. CoQ10 = coenzyme Q10; Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg.

more important than the Q100 group in driving the significance of the relationship linking CoQ10 level change to change in SPS (and trend to change in GSRH).

An exploratory analysis assessed potential foundations for the more favorable finding on GSRH with Q100 in men versus when women were added. CoQ10 levels increased, and achieved CoQ10 levels in women on Q100 were less than in men: for instance, mean CoQ10 levels on Q100 in men were 3.0 (range 2.3–4.1), while the mean value of 2.3 in women matched the lowest achieved level in men (range 1.5–3.3). The GSRH change was also less favorable. Differences in GSRH response profiles for women versus men, on both placebo and Q100 (perhaps in part due to CoQ10 change differences relative to men) led to magnified variance and SE for GSRH change on both placebo and on Q100 when women were included, attenuating power for a comparison. Of note, in the (sole) woman on Q300, the CoQ10 level increased twice as much as the average increase in men (+3.7 in women versus +1.8 in men), and the associated GSRH change was also strikingly more favorable (difference versus men +1.28). These factors may contribute to focus of the GSRH benefit in the male subsample, while women nonetheless contribute to the CoQ10-change relationships to the outcome.

GSRH and symptoms are both self-rated, and since there were GSRH disparities at baseline, we evaluated if baseline GSRH disparities affected symptom change scores (see section 2). Higher baseline GSRH significantly predicted greater reporting of symptom benefit on placebo, and a significant treatment × baseline GSRH interaction was observed for many symptoms, necessitating adjustment for the baseline GSRH × treatment interaction (and its components). The small sample poorly supported the number of adjustments. We emphasize low confidence in the actual values for each
## Table 5: Relation of Change in CoQ10 Level to Change in Self-Rated Health and Function

<table>
<thead>
<tr>
<th></th>
<th>All Participants</th>
<th>Excluding Q300 Arm</th>
<th>Excluding Q100 Arm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>β (SE)</td>
<td>95% CI</td>
</tr>
<tr>
<td>GSRH</td>
<td>36</td>
<td>.15 (.08)</td>
<td>-.01, .32</td>
</tr>
<tr>
<td>SPS</td>
<td>32</td>
<td>.31 (.14)</td>
<td>.03, .59</td>
</tr>
</tbody>
</table>

Notes: β = Regression coefficient; CI = 95% confidence interval; GSRH = general self-rated health; N = number of participants in analysis; SPS = summary performance score. Analyses adjust for baseline values of the variable and baseline CoQ10 level. Participants who were at the ceiling for the outcome at baseline (in each randomization arm) were excluded from analysis. Placebo participants also contribute and are appropriate to include, as they add both some participants with negative changes in CoQ10 level (to be assessed against outcome) and also some materially positive changes in CoQ10 level.
symptom and provide the results qualitatively only. The sign of the regression coefficient for Q100 versus placebo was favorable for 19 of the 20 symptoms, $p = 0.00004$ (sign test), though not all differences were clinically relevant. The sole symptom for which the direction was not favorable was sleep problems (the effect was nonsignificant, but the direction was not favorable), in the setting of known potential for activation insomnia with CoQ10. Benefits and trends emphasized those with low GSRH, with attenuation of benefit with higher GSRH (consistent with the placebo benefit finding). Symptoms for which favorable Q100 regression coefficients reached significance included trouble recalling words or names, impatience, irritability, headache, fatigue with exertion, problems with energy to do things, and muscle pain. Trends to benefit were observed for need to go back and recheck things ($p < 0.1$) and for joint pain ($p = 0.10$). Considering only symptoms for which $p \leq 0.1$, the sign test $p$ was 0.0039 (two-sided), still strongly significant.

In contrast, Q300 bore no consistent direction of effect relative to placebo. However, significant favorable effects were seen for fatigue with exertion, sleep problems, several cognitive variables (attention or concentration, memory tasks, trouble recalling names or words), cold hands or feet, and tinnitus. Unfavorable effects were observed for aches and pains, and the correlated muscle pains, and for energy to do things. We reiterate that adjustment for the GSRH interaction (and its component terms) was requisite for conceptual validity, but the sample poorly supported the number of adjustments, so findings for individual symptoms should be viewed as highly provisional.

No effects on digit span backward were observed. Digit span backward was predicted to have poor sensitivity relative to cognitive self-ratings. Cognitive tests show commonly large training effects with high variance, typically requiring larger samples to see even sizable effects, and single measurements are disadvantaged relative to subjective ratings, which intrinsically average performance over a time window. Indeed effects were nonsignificant. However, measurement of digit span did afford objective validation of cognitive symptom self-ratings. Digit span (forward) correlated to assessed self-ratings: attention or concentration problems, $r = -0.33$, $p = 0.049$, trouble recalling words or names, $r = 0.35$, $p = 0.023$, reading comprehension, $r = -0.41$, $p = 0.018$, and trouble remembering where going or what doing, $r = -0.49$, $p = 0.0035$. No effects on visual analog scales of pain intensity or pain unpleasantness were observed.

Two significant adverse events were reported during the study. One 40-year-old male veteran experienced a neurological event (represented as a possible seizure but with a description consistent with a transient ischemic attack) at about 3 months into study participation while visiting family in a distant state. His physician conjectured the cause might be an interaction between two prescribed CNS-active drugs, one of which had been recently initiated. That medication was discontinued, with no recurrence. (Unblinding
at study conclusion showed he had been on placebo.) One 52-year-old female veteran experienced a stroke, also while traveling out of state, that led to right-sided weakness and expressive aphasia, which improved with time based on follow-up calls (she did not continue participation). However, she was off all study treatment at the time (she had been intended to be on placebo but had taken nothing for a month prior to the event). Study participation was considered unlikely to relate to either event: CoQ10 would be “expected” to confer relative protection from seizure or stroke (Grieb, Ryba, Sawicki, & Chrapusta, 1997; Lalani et al., 2005; Ogawa, Tsukamoto, Hirose, & Kuroda, 1986) (with human data supporting benefit to other vascular events) (Singh et al., 1998), but in any case, both participants were assigned to placebo at the time of their events.

Regarding milder adverse events, from literature and experience with CoQ10, we anticipated occurrence of activation insomnia with evening use for some, and a contingency plan was in place prior to study initiation by which participants who called with sleep problems (without unblinding) were advised to take their evening study treatment earlier. No communications by participants of other (nonsleep-related) adverse events occurred.

4 Discussion

Q100, in a high-quality/bioavailability CoQ10 preparation (PharmaNord, Denmark) and compared to placebo in a small pilot study, significantly benefited GSRH, the primary outcome, in men, representing 85% of participants, but not in the total sample. Men constitute most GWV, rendering this an important group, but findings should not be overinterpreted. Physical function, assessed by SPS, was significantly favorably affected by Q100 in the full sample (females contributed). Perhaps most important, greater change in CoQ10 levels in the total sample (including a Q300 group) significantly predicted change in physical function and were near-significant in predicting change in GSRH.

For symptoms, favorable direction coefficients with Q100 were present for 19 of 20 symptoms. Nine symptoms showed statistically significant or borderline significant differences ($p \leq 0.1$) comparing Q100 versus placebo, with each favoring Q100. For Q300, no consistent direction of effect relative to placebo was observed. Several significant favorable symptom effects were noted. These included exertional fatigue as well as several cognitive symptoms, in common with Q100. However, some symptoms were unfavorably affected: two related to muscle pain and one to energy to do things. One hypothesis could be that reduced fatigue with exertion may have led to greater exertion, prior to an opportunity for adequate muscle recovery (time is required for muscle satellite cells to restore muscle cells that, by hypothesis, may have been lost during periods of impaired energy), producing pain, which may inhibit action (doing of things).
Although benefit was preferentially observed with Q100, the relationship of outcome change to CoQ10 level change, excluding those at the outcome ceiling at baseline, actually lost more significance when the Q300 than when the Q100 group was excluded. It is important to note that the analysis by CoQ10 level incorporates the reality of CoQ10 change, even where this departs from the intent, including any placebo participants who may have taken CoQ10 outside of that intended in the study (one placebo participant had a greater CoQ10 rise than any on Q100), and any whose CoQ10 level failed to rise despite allocation to a CoQ10 treatment arm. The basis for little rise in CoQ10 levels for some participants assigned to Q100 or Q300 is not clear. Noncompliance is a possibility: participants are veterans with CNS problems, and both cognitive and mood problems have been shown to predict medication noncompliance (Stilley, Sereika, Muldoon, Ryan, & Dunbar-Jacob, 2004). Barriers to assimilation may be present in some; gastrointestinal problems such as diarrhea are not uncommon in GWI, and pancreatic exocrine dysfunction can accompany mitochondrial problems (Hsu et al., 2005), which is reported in GWI (Koslik, Hamilton, & Golomb, 2014). At least one participant acknowledged having disregarded instructions to take one softgel from the smaller bottle or allotment and two daily from the larger, mixing all together. For any such participants randomized to Q100 (only), CoQ10 levels measured at a point in time need not reflect levels over the recent course of that treatment, producing more dissociation between measured level and assessed health. Despite any such factors, in the total sample, a relation of CoQ10 level change to SPS change is significant, and a relation of CoQ10 level change to GSRH change approached significance, adding strong support to a true, causal benefit from CoQ10 supplementation. The Q300 group was more important than the Q100 group in establishing this relationship.

Significance for clinical improvement as a function of CoQ10 level depends on inclusion of the Q300 (but not Q100) group, yet estimated magnitude per CoQ10 change was greater for the Q100 than Q300 group, and significance was greater overall for the Q100-placebo comparison. Findings from the study help reconcile these findings.

First, some on Q300 did not change their CoQ10 levels much, nor did their outcomes change much. These individuals diminish the assessed average effect of Q300 (versus placebo) on the outcome but contribute to the relation of CoQ10 change to outcome change in the Q300 arm.

Second, some on Q300 did increase CoQ10 levels but had no ability to improve the outcome, since they were already at the ceiling on the outcome at baseline. (This disproportionately affected the Q300 group.) These individuals diminish the assessed average effect of Q300, in the intention-to-treat analysis, but do not adversely affect the analysis designed to understand how CoQ10 change related to outcome change, in which those already at the ceiling at baseline were excluded from analysis (for all randomization groups) as they could not in principle improve. (There were no floor
effects.) Those near the ceiling were not excluded and will have precluded a larger rise in CoQ10 level on Q300, producing a proportionately larger improvement in the outcome (as is necessary for the coefficient to retain the magnitude seen in the sample with Q100). (Those who do have some room to improve can contribute to instances of larger CoQ10 level changes on Q300 linked to larger outcome change, producing greater statistical leverage, and providing for Q300 disproportionately driving the significance of the relation of CoQ10 level change to outcome change.)

Third, some on placebo showed an increase in CoQ10 level despite nominally not taking CoQ10. When such a rise in level is tied to an improvement in outcome, it attenuates significance for the Q300 (and also Q100) versus placebo comparisons, but contributes to the CoQ10-change versus outcome-change assessment.

Revisiting the second point, the Q300 group had higher baseline scores on GSRH and SPS than the other randomization groups; room for improvement is curtailed or nonexistent for participants already near or at the ceiling, respectively. The Q300 group, for instance, had the sole study participant who rated GSRH at the ceiling of 5 ("excellent") at baseline, in whom no GSRH improvement on treatment was possible. Additionally, 4 of 12 (a third) who were randomized to the Q300 group were at the SPS ceiling at baseline—two or more times the fraction who were at the ceiling in either the placebo or Q100 groups. These participants who could not improve—more of whom were in the Q300 arm—were included in the primary intent-to-treat analysis. Their inclusion spuriously depresses coefficients and significance for any Q300 benefits. These participants were excluded in the analysis, which seeks to understand the relation of CoQ10 change to outcome change.

As a further factor differentiating the lower- and higher-dose groups, which could diminish observed effect with the higher dose, nighttime CoQ10 administration is systematically present in the Q300 group only and may result in sleep problems for some that may adversely affect outcomes. A larger study, incorporating Q300 without (and with) a nighttime dose, could determine whether and to what degree this nighttime use influenced outcomes.

To our knowledge, this is the first randomized, double-blind treatment trial to report benefit of a treatment in GWV with GWI only. Previously, the few treatments supported with randomization in GWI have involved therapies where blinding was not feasible; that are time, practitioner, and thus cost intensive, with feasibility and compliance obstacles (cognitive-behavioral therapy and exercise; Donta et al., 2003); or required the presence of other diagnosed conditions that would be expected to benefit from the treatment with or without GWI (e.g., sleep-disordered breathing was required for eligibility, for a trial of CPAP (Amin et al., 2011), and CPAP has been reported to benefit numerous symptoms in those with sleep-disordered breathing, in the absence of GWI (Barnes et al., 2002;
Flemons & Tsai, 1997; Onen et al., 2010; Shepherd, Hillman, Holloway, & Eastwood, 2011; Weaver et al., 2012). A prior small pilot study of carnosine as an antioxidant was reported as generally negative (albeit with a sample size of 25 participants); however, possible benefits to cognitive indices and diarrhea were reported (Baraniuk, El-Amin, Corey, Rayhan, & Timbol, 2013). Diarrhea was not among symptoms present in more than 50% of our participants. (In a post hoc analysis, diarrhea showed a borderline-significant favorable-direction relation to CoQ10 level change scores, despite the smaller sample ($p = 0.088$).) An emphasis on cognitive benefits for both CoQ10 doses is consistent with results from our study. CoQ10, beyond its antioxidant effects, more directly supports cell energy production.

As noted above, brain function, including neural processing, is heavily energy demanding (Erecinska & Silver, 2001; Fehm et al., 2006; Shulman et al., 2004); so too is muscular exertion. We suspect it is not coincidental that exertional fatigue and cognitive symptoms were more consistently favorably affected symptom domains, benefiting with both CoQ10 doses.

The study was limited by a small sample size, which restricts power to see effects. Power was further constrained by participation of individuals at and near the ceiling of measures assessed. Nonetheless, there was a significant relation of change in CoQ10 level to change in physical function and a near-significant relation to change in GSRH, supporting a causal benefit of CoQ10 to function and adding the likelihood that this benefit will extend to general health in a larger sample. Order, counterdirectional, and carry-over-related effects obviated the validity of an originally intended crossover approach, directing focus to the randomized, parallel design study embodied in the first phase, that retains validity in this setting. The lesser power associated with the parallel design underscores the large apparent effect sizes where significance was observed, and a relation of benefit to CoQ10 change provides strong evidence to suggest causality. There were disparities in baseline GSRH, requiring adjustment for baseline values. Participant self-selection could underrepresent those most and least severely impaired. (Clinically, omission of those most affected is the greater concern.) Authority of findings will rest on replication.

This study provides valuable information for future studies concerning dosage, design considerations and desiderata (see also appendix B), and outcome measures for evaluating the effects of this supplement.

Regarding lessons learned, there are several. Key outcomes included GSRH and SPS, but inclusion was based solely on adherence to GWI criteria, which is based on symptoms: some participants were already at or near the ceiling of these measures and could not contribute to power for these outcomes, a problem in a small study. This could be addressed by a larger study (with a power buffer), by avoiding participants at or near the ceiling of the measures of primary interest, or by modifying the measures to allow more grain, perhaps with a more continuous outcome.
Related to this, a general health measure with a finer grain may enhance sensitivity (and also reduce risk of ceiling effects). A participant who perceives his or her health to transition from barely more than “fair” to barely less than “very good” may have had a meaningful change but in both cases will represent his or her health as “good.”

Third, we would implement approaches to better address regression-dilution bias and general variance. One approach might include a longer run-in, or selection of participants who previously met inclusion criteria (with affirmation that they still do). This is because participants who met symptom thresholds for symptoms that vary in time are expected on average to regress, variably, to the mean; where possible, this should be permitted to occur at least in part prior to the baseline visit. Another approach would be repetition of assessments at baseline and then again on treatment, averaging assessments at each time point to increase the precision of estimates, thereby enhancing power at a given sample size. This may be important for subjective assessments where anchoring may add variance and also for measures like cognitive tests, which not only have intrinsic test-retest variance but individual differences in training effects. To illustrate the benefit to power of repeat assessment, in this study self-rated muscle weakness did not significantly relate to measured physical function by SPS at either the run-in visit ($r = -0.20, p = 0.18$) or the baseline visit ($r = -0.28, p = 0.053$; note that significance is already improved, suggesting some of the variance has been addressed by training or other forces, even for these noncognitive assessments). However, a significant relation between these measures was apparent when the run-in and baseline scores were averaged for each of the two measures ($r = -0.34, p = 0.024$).

Fourth, we would employ absolute ratings for symptoms in preference to (or in addition to) change ratings. Participant burden forced us to choose one of the two, and we had prior validity data only for symptom change ratings (in a different setting). (See further discussion in appendix B.)

In summary, in this study, Q100 benefited physical function and symptoms in GWV with GWI and benefited GSRH in the male subset. In the total sample, which also included a Q300 arm, CoQ10 level changes significantly positively predicted physical function change and were near significant in predicting change in GSRH. These relations to CoQ10 levels suggest a causal benefit of CoQ10 in GWI in the presence of a favorable overall risk-benefit balance. Findings warrant replication and extension, with careful methods and retaining a high-quality CoQ10 preparation, in a larger sample.

Appendix A: Eligibility

Eligibility required that subjects meet the following conditions:

1. Deployment criteria: Deployed to the Middle East at any time between August 1990 and July 1991.
2. CDC criteria for Gulf War illness (Fukuda et al., 1998): Symptoms present for more than 6 months in at least two of three categories: fatigue, musculoskeletal, mood/cognition.

3. Kansas criteria for Gulf War illness: Multiple symptoms within the category or symptoms of at least moderate severity (not “mild”) in each of at least three of the following six categories: pain, respiratory, fatigue/sleep, gastrointestinal, neurological/cognitive/mood, and dermatologic. To meet Gulf War illness criteria, symptoms must have persisted or recurred in the year prior to interview and first have been a problem for respondents in 1990 or later (Steele, 2000).

4. Exclusion criteria: Subjects could not have other medical diagnoses that can cause these symptoms (e.g., lupus, multiple sclerosis).

Appendix B: Cross-Over Versus Parallel Design

The study was originally conceived with a cross-over design. Several factors conspired to render a cross-over analysis uninterpretable. Considerations related to nonmonotonic effects in the setting of use of change scores, carry-over effects, ceiling effects, and order effects. We characterize these in more detail below. Together these factors obviated validity of any cross-over analysis and led the primary analysis to be revised, to focus on the (first phase) parallel design randomized trial, which retains full validity.

B.1 Method. All subjects received placebo during two of four 3.5 ± 0.5-month treatment phases and coenzyme Q10 (CoQ10) during the remaining two phases—one at higher dose (Q300) and one at lower dose (Q100). Placebo phases were designed to also serve as a washout between CoQ10 phases. Either Q100, Q300, or placebo could come first (placebo in about twice as many subjects in each phase, to provide for Q100 and Q300 in the next and/or prior phase), providing four orders.

B.2 Considerations. The initial plan had been to include both absolute and change scores for symptoms at each visit, but feedback during protocol testing indicated that participant burden was excessive and we were able to retain only one, except we kept absolute ratings at baseline (to adjust for these) and the final visit (at which participant burden would not compromise continued study participation). The choice to retain change scores was based on the fact that these had been prespecified for symptoms because we had experience and validation data with these from another study and under the presumption that these might enhance power by preserving at least the directionality of change, for participants in whom anchoring and other effects led to renorming. This was recognized to have risks in a crossover design in the event of nonmonotonicity. Greater benefit with the lower dose is recognized to be a real or perhaps expected possibility.
for nutrient or antioxidant treatment, depending on positioning within the optimal dose range (Miller et al., 2005; Morris & Tangney, 2011).

B.2.1 Change Score Effects. In a single-dose crossover study, change scores on and off an agent (the latter “on” placebo) can be compared, and are expected to show the opposite sign, enhancing power. However, nonmonotonic effects of Q300 versus Q100 meant that change scores on placebo and on Q100 bore the same sign. (Change score ratings also had a coarse grain: symptoms designated as only either somewhat or much greater, or less, do not permit discrimination between a modest improvement going off Q300, onto placebo, and a moderate improvement going onto Q100, hypothetically speaking.) This change score issue compromises validity of a crossover, but does not compromise validity of the first-phase randomized parallel design assessment.

B.2.2 Carryover Effects. True carryover effects in CoQ10 levels were also present, though these were most evident at the final phase and not equally evident in all groups. In those who had undergone treatment with CoQ10, placebo periods did not consistently restore CoQ10 levels to previous levels, consistent with carryover effects. In consequence, following the final placebo phase, mean CoQ10 values were significantly higher than CoQ10 values at baseline, particularly for the group coming off Q300, consistent with carryover effects for CoQ10 \((p < 0.001)\).

The CoQ10 carryover effects alone are not the primary factor that led to loss of validity for the carryover, though at the conclusion of the final phase, it is clear these had an impact. Table 6 shows that CoQ10 level differences were strongly significant between the Q100 group and the placebo group at the conclusion of the first phase but not at conclusion of the last phase. Table 7 shows the treatments at each phase by randomization group. Figure 4 shows also that CoQ10 levels on placebo at the final visit (in the group treated with Q300 in the prior phase) were not only higher than at baseline but were higher than on Q100 at that final visit. This figure depicts graphically the finding corresponding to Table 6 that there was no significant difference between CoQ10 levels on Q100 versus placebo following the final visit.

B.2.3 Ceiling Effects. Related to carryover effects and possibly training effects, more participants were at the ceiling, without the ability to improve, at later visits. Without the ability to improve, comparison to an earlier-phase treatment in which there was ability to improve is not interpretable. This issue was minimized with a focus on the first-phase parallel design analysis. (For instance, at baseline, 21.7% were at the Summary Performance Score ceiling. This increased to 45% by conclusion of the first treatment phase and was still greater, 50% or more, at later phases.)
Table 6: Differences in Coenzyme Q10 levels ($\mu$m/ml) Between Q100 and Placebo Groups.

<table>
<thead>
<tr>
<th></th>
<th>Q100</th>
<th>Placebo</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>95% CI</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>CoQ10 level, end of first phase</td>
<td>2.81 (0.65)</td>
<td>2.34, 3.27</td>
<td>&gt; 1.65 (0.70)</td>
</tr>
<tr>
<td>CoQ10 level, end of last phase</td>
<td>2.23 (1.01)</td>
<td>1.51, 2.96</td>
<td>&gt; 1.97 (0.91)</td>
</tr>
</tbody>
</table>

Notes: CI = 95% confidence interval; CoQ10 = coenzyme Q10. The differences were strong for the first phase and nonsignificant following the final phase.
Table 7: Treatments by Randomization Group and Phase.

<table>
<thead>
<tr>
<th>Group</th>
<th>Phase 1 (Month 0–3.5)</th>
<th>Phase 2 (Month 3.5–7)</th>
<th>Phase 3 (Month 7–10.5)</th>
<th>Phase 4 (Month 10.5–14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Placebo</td>
<td>Q100</td>
<td>Placebo</td>
<td>Q300</td>
</tr>
<tr>
<td>Group 2</td>
<td>Q100</td>
<td>Placebo</td>
<td>Q300</td>
<td>Placebo</td>
</tr>
<tr>
<td>Group 3</td>
<td>Q300</td>
<td>Placebo</td>
<td>Q100</td>
<td>Placebo</td>
</tr>
<tr>
<td>Group 4</td>
<td>Placebo</td>
<td>Q300</td>
<td>Placebo</td>
<td>Q100</td>
</tr>
</tbody>
</table>

Note: Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg.

Figure 4: Coenzyme Q10 level change with treatment. Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg. m3, m6, m9, m12 are simplifications to the end of the respective 3.5 ± 0.5 month treatment periods.

B.2.4 Order Effects. Order effects were present, related to carryover effects, to ceiling effects, and possibly to the use of a measure involving thresholding in baseline selection (a setting in which there is regression to the mean, which may be more pronounced in earlier phases). This poses a problem for a crossover comparison, in which the order of the compared treatments
Table 8: Dropouts by Phase.

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Q100</th>
<th>Q300</th>
<th>Total Dropouts by Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropouts in phase 1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Dropouts in phase 2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Dropouts in phase 3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Dropouts in phase 4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total dropouts across phases</strong></td>
<td><strong>5</strong></td>
<td><strong>3</strong></td>
<td><strong>4</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>

Notes: Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg. Final total: 46 − 12 = 34 (of goal completer number 32).

is necessarily different. This does not compromise the validity of the first-phase randomized parallel design assessment.

Together, these factors necessitated revision of the primary analysis to focus on the parallel-design first-phase analysis. Use of the first-phase parallel design also has the advantage that participant attrition over the course of the study works more strongly to compromise comparison in the crossover approach: where a subject was missing for either phase of the comparison, the comparison cannot be undertaken. For simplicity of exposition, we refer to the composite set of issues that compromised the crossover approach as carryover effects.

**B.3 Retention.** Analysis was refocused on the first-phase parallel analysis, as described below. For that reason, the consort figure (see Figure 2) provides dropout values through that phase only to reduce confusion. Table 8 shows dropouts throughout the full study.

**B.4 Power Considerations.** The initial aim was to recruit sufficient participants to achieve 32 completers of the full crossover study, understanding that the special considerations and hardships that attend Gulf War illness would lead to challenges and some expected attrition (which, indeed, occurs in all studies). This aim was achieved. With 46 participants recruited, 42 completed the first phase, and 34 completed the study. The power calculations presumed 32 available for comparison to themselves in a crossover design. The revision to the parallel design approach provides an apparent larger number, for completers of the first phase, but a smaller number per group, and somewhat attenuated power.

Appendix C: Details of Measures Assessed: Modifications in Measures and Analysis

**C.1 Single-Item General Self-Rated Health** (DeSalvo, Bloser, Reynolds, He, & Muntner, 2006; DeSalvo, Fan, McDonell, & Fihn, 2005; DeSalvo, Fisher
et al., 2006). At baseline subjects were asked, “In general, for the last two weeks, would you say your health is poor, fair, good, very good, or excellent?” On follow-up they were asked, “Since your last visit, in general, for the last two weeks, would you say your health is poor, fair, good, very good, or excellent?” This assessment has coarse grade. It shows predictive validity in large samples and was chosen as the primary outcome because of ease of administration and a citable literature on psychometric properties.

C.2 Physical Function: Summary Performance Score. This is the sum of scores for three items assessing physical function (Guralnik et al., 1994) that has been reported to have high reliability and good sensitivity to change in settings where it has been assessed (Ostir, Volpato, Fried, Chaves, & Guralnik, 2002). The assessed items are time to complete five chair rises, timed standing balance three ways, and 4 meter walking velocity. Each item is scored 0 to 4, providing a maximum score of 12 (higher score = higher function). SPS has been shown to significantly predict incident disability (Guralnik et al., 2000).

According to McDermott et al. (2006):

Repeated Chair Rises. Participants sat in a straight-backed chair with their arms folded across their chest and rose to a standing position, repeating the exercise 5 consecutive times as quickly as possible. We measured the time each patient required to complete 5 chair rises.

Standing Balance. Participants were asked to hold 3 increasingly difficult standing positions for 10 seconds each: standing with feet together and parallel (side-by-side stand); standing with feet parallel, with the toes of 1 foot adjacent to and touching the heel of the opposite foot (semi-tandem stand); and standing with 1 foot directly in front of the other (tandem stand).

Walking Velocity. Walking velocity was measured with a 4-meter walk performed at “usual” and “fastest” pace. For the usual-paced walk, participants were instructed to walk at their usual pace, “as if going down the street to the store.” For the fastest-paced walk, participants were instructed to walk as fast as they could. Each walk was demonstrated by the research assistant. Participants were given the command “ready, go”; timing began on “go.” Each walk was performed twice, and the faster time in each pair was used in analyses.

Summary Performance Score. The summary performance score combined data from the usual-paced 4-meter walking velocity, time to rise from a seated position 5 times, and standing balance. Individuals received a score of 0 for each task they were unable to complete. Scores ranging from 1 to 4 were assigned for all completed tasks; the scoring system was based on quartiles of performance for over 5000 participants in the Established Populations for the Epidemiologic Study of the Elderly.
Scores were then summed to obtain the summary performance score, which ranged from 0 to 12.

**C.3 CoQ10 Measurement.** Analyses were performed by Dennis Health in Cheryl Rock’s Nutrition Research Laboratory at UCSD. Serum samples were collected, separated, and frozen at −80°C until analyzed. High-performance liquid chromatography (HPLC) was used to separate and quantify CoQ10 (equipment from Varian, Walnut Creek, CA: a model 410 auto sampler, 325 UV/VIS dual wavelength detector, and Prostar 230 reagent pump). The Complete CoQ10 HPLC Kit (ALPCO Immunoassays, Salem, NH) was used for the two-step procedure: rapid extraction of CoQ10 from serum with an organic-based precipitant and internal standard, then centrifugation to remove protein precipitants. The upper organic phase of the treated sample is removed and evaporated. The residue is resuspended in ethanol and an aliquot assayed on a Varian HPLC system with Starworks software by injection onto a reverse-phase column heated at 30°C. Separation and detection is at 275 nm, and calculation is based on peak height. One calibrator and five quality control samples are analyzed with each batch of samples. Method performance data show linearity up to 10 µg/mL, with a lower detection limit of 0.02 µg/mL. The intra-assay consistency value is 4.4% at 0.66 µg/mL; 6.6% at 0.31 µg/mL, and 4.5% at 0.89 µg/mL. To monitor HPLC method performance, we participate in the proficiency survey of the National Institute of Standards and Technology (NIST) for fat-soluble vitamins (including CoQ10), which involves three round-robin blind analyses of plasma samples provided by NIST each year. We routinely used one in-house serum pool and two additional purchased quality control (QC) samples. The batched sample results were accepted only if these internal QC results were within two standard deviations of the assigned values.

**C.4 Individual Symptoms.** Individual symptoms, among those expressed by at least half of subjects, were examined.

**C.5 Use of Change Ratings for Gulf War Illness Symptoms.** As above, subject visit burden in pilot testing led us to choose between absolute self-ratings or self-ratings of change in symptoms on follow-up (versus assessing both). Change scores were used for follow-up assessment, because these scores were prespecified. This was on the reasoning that absolute ratings are subject to anchoring and other recalibration effects (Wilson, Houston, Etling, & Brekke, 1996), which can lead the “same” level of symptom to receive different ratings on different days, or different symptoms to receive the same rating, or ratings whose difference bears the wrong sign. Our reasoning was that directly rating the change reduced the risk of subverting the direction of difference, with the potential to improve power.
Change ratings can confer benefit in the setting of classical carryover effects but hold a risk in recurrent crossovers if opposing direction effects occur for different doses. Such change ratings have shown good psychometric properties and sensitivity to change (Clifford et al., 2005; Golomb, Broadwin, White, Criqui, & Dimsdale, 2009; Golomb, Kwon, Criqui, & Dimsdale, 2007). Change scores had been effective and sensitive in a prior study (rating symptoms on follow-up from “much worse” to “much better” than the baseline of that treatment phase as here) and showed good convergent validity against validated instruments, leading us to select change scores for this study. However, that prior study was much larger, perhaps better accommodating the coarse grain of the change score (no distinction between small perceived changes and moderate changes).

Absolute scores were included at the baseline visit (to permit adjustment for potential regression to the mean) and the final visit only (where the additional burden would not compromise participant retention). Change scores can confer benefit to power in the setting of the more commonly recognized type of carryover effect, benefiting analysis provided there is a consistent direction of change for both active treatment arms relative to placebo. However they pose a risk of introducing a different type of carryover effect if nonmonotonic effects of the active treatment options are present. In this case, benefit can arise with transition to placebo—from a putatively relatively unfavorable higher dose, which can erode effect size and thus power to detect benefit with transition (from placebo) to the lower dose.

Proposed study outcomes are shown in Tables 9 and 10.

C.6 Revision to First-Phase Parallel Design Analysis: Reasons

C.6.1 Carryover Effects. CoQ10 levels following the final postplacebo phase (in subjects who had now received 3.5 month treatment periods on Q100 as well as on Q300) differed significantly from baseline CoQ10 levels: CoQ10 (mean ± SD) after placebo in the final phase: 1.97 ± 0.91, versus pretreatment averaged CoQ10 1.48 ± 0.48, difference 0.49 (SE 0.11) 95% CI 0.28-0.72, p = 0.0001. (Relative to within-phase CoQ10-level comparisons across randomization arms, power for this comparison benefits from larger sample size, including use of the full sample at baseline.) CoQ10 levels following assignment to placebo in treatment period 1 (no prior CoQ10 treatment period) did not differ from CoQ10 levels at baseline.

C.6.2 Blunted Final Phase CoQ10 Differences Across Treatment Assignment. CoQ10 values following treatments in the final phase bore the expected direction of differences by treatment assignment, but differences were blunted, SDs were higher for placebo and Q100 groups than after the first
Table 9: Proposed Study Outcomes (Primary and Secondary).

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Prespecification</th>
<th>Proposed</th>
<th>Conducted</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall health impact</td>
<td>Primary, specific aim</td>
<td>GSRH</td>
<td>GSRH</td>
<td>Reliability/validity not strong. Near significant ((p \sim 0.05)) lack of baseline comparability.</td>
</tr>
<tr>
<td>Muscle function</td>
<td>Secondary, specific aim</td>
<td>Summary performance score</td>
<td>Summary performance score</td>
<td></td>
</tr>
<tr>
<td>CoQ10 levels (objective)</td>
<td>Secondary, specific aim</td>
<td>CoQ10 levels, total and reduced</td>
<td>CoQ10 levels, total only</td>
<td>Reduced CoQ10 levels must be done at the site of collection due to rapid oxidation ex vivo. This was not possible under funds available.</td>
</tr>
<tr>
<td>CoQ10 effect modification</td>
<td>Secondary, specific aim</td>
<td>CoQ10 effect modification</td>
<td>CoQ10 effect modification</td>
<td></td>
</tr>
<tr>
<td>Dose comparison (secondary, specific aim)</td>
<td>Secondary, specific aim</td>
<td>Dose comparison</td>
<td>Dose comparison</td>
<td></td>
</tr>
<tr>
<td>Individual symptoms (secondary, specific aim)</td>
<td>Secondary, specific aim</td>
<td>Self-rating</td>
<td>Self-rating</td>
<td></td>
</tr>
<tr>
<td>Oxidative stress</td>
<td>Secondary, specific aim</td>
<td>MDA</td>
<td>Samples were procured and sent to a specialty lab but error (an apparent result of “sorting” of EXCEL columns by the test laboratory, without retention of the original values in a separate format) led to loss of information connecting values to subject IDs and visits. Data rendered unusable.</td>
<td>Unable to assess. Only four subjects had valid baseline and on-treatment values (a small uncompromised first database had some subjects on baseline, but only four had baseline and any follow-up value). For these four, Q100 had a more favorable effect on MDA than did Q300 or placebo, consistent with a possible role for change in oxidative stress as a mediator (but the usable sample was far too small for conclusions).</td>
</tr>
</tbody>
</table>
Table 9: Continued.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Prespecification</th>
<th>Proposed</th>
<th>Conducted</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognition</td>
<td>Secondary, not in specific aims</td>
<td>Self-rating (4 symptoms within the symptom score)</td>
<td>Self-rating (4 symptoms within the symptom score)</td>
<td>Self-ratings were characterized, in advance, as likely to have greater sensitivity than “objective” measures, which suffer from variance due to variable training effects (among other issues). Self-ratings were validated against an objective measure in this study. As prespecified, subjective assessments “in a blinded setting, will enable subjects to tap the elements of subjective assessment that have been shown to have greater sensitivity than neuropsychological tests, while blinding and randomization preclude opportunity for subjective estimates to be influenced by knowledge of treatment assignment.”</td>
</tr>
<tr>
<td>Pain (character)</td>
<td>Secondary, not in specific aims</td>
<td>Pain intensity, pain unpleasantness</td>
<td>Pain intensity, Pain unpleasantness</td>
<td>No effect on pain character—reported intensity or unpleasantness.</td>
</tr>
</tbody>
</table>

Note: CoQ10 = coenzyme Q10; GSRH = general self-rated health; MDA = malondialdehyde; Q100 = coenzyme Q10 100 mg; Q300 = coenzyme Q10 300 mg.
phase. This, together with some sample attrition, contributed to loss of statistical significance of CoQ10 differences across treatment arms (placebo 1.97 (SD = 0.91), Q100 2.23 (SD = 1.01) \( p = 0.49 \) versus placebo, Q300 2.49 (SD = 1.91) \( p = 0.39 \) versus placebo).

Table 10: Additional Secondary Proposed But Not Measured.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Test</th>
<th>Reason Not Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other cognition</td>
<td>Dual function</td>
<td>No specific dual function test identified. Not conducted.</td>
</tr>
<tr>
<td>Other muscle</td>
<td>Nicholas manual muscle test, wrist actigraphy (activity)</td>
<td>Not conducted.(^a) Time burden, cost.</td>
</tr>
<tr>
<td>Other fatigue</td>
<td>MFSI</td>
<td>Time burden, duplication.</td>
</tr>
</tbody>
</table>

\(^a\)Subject visit/time burden (and in some instances cost), guided by study priorities, led to nonimplementation of some proposed measures. MFSI = multidimensional fatigue symptom inventory.

C.6.3 Counterdirectional Symptom Effects of Q300 versus Q100. These are discussed in appendix B.

Appendix D: Baseline Self-Ratings of Health Prior to the Gulf War (Retrospective) and Now (Contemporaneous)

Table 11: Self-Rated Health Status Before the Gulf and at Run-In (%).

<table>
<thead>
<tr>
<th></th>
<th>Excellent</th>
<th>Very Good</th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health before Gulf (retrospectively assessed)</td>
<td>74</td>
<td>22</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Health now</td>
<td>2</td>
<td>4</td>
<td>22</td>
<td>57</td>
<td>15</td>
</tr>
</tbody>
</table>

Acknowledgments

This study was funded by the Department of Defense Congressionally Directed Medical Research Program GW060036. PharmaNord (Denmark) provided Coenzyme CoQ10 and matching placebo softgels for this study. We sincerely thank the Gulf War veterans, who generously gave their time to this effort, and the study personnel without whom this would not have been possible. We thank Cheryl L. Rock, Dennis Heath, and Julie O. Denenberg for their assistance with the study and comments on the manuscript. Conflict of interest: No competing interests exist. Funding statement: The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the letter. Comment: Findings from this study...
were presented publicly (Golomb, 2011). Specifics of some findings differ due to refinements to the analysis approach.

**References**


CoQ10 and Gulf War Illness 2639


Lee, B. J., Tseng, Y. F., Yen, C. H., & Lin, P. T. (2013). Effects of coenzyme Q10 supplementation (300 mg/day) on antioxidation and anti-inflammation in coronary


Spat, A., Fulop, L., & Szanda, G. (2012). The role of mitochondrial Ca(2+) and NAD(P)H in the control of aldosterone secretion. *Cell Calcium*, 52(1), 64–72.


hypopnea syndrome. Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi, 43(3), 179–182.

Received August 9, 2013; accepted May 28, 2014.