Author’s response to reviews

Title: A review of epidemiologic research assessments of United States Gulf War I era veterans

Authors:

Rebecca B McNeil (rebecca.mcneil@va.gov)
Catherine M Thomas (catherine.thomas4@va.gov)
Steven S Coughlin (steven.coughlin@va.gov)
Elizabeth Hauser (elizabeth.hauser@va.gov)
Grant D Huang (grant.huang@va.gov)
Karen M Goldstein (karen.goldstein@va.gov)
Marcus R Johnson (marcus.johnson4@va.gov)
Tyra Dunn-Thomas (tdunnthomas@gmail.com)
Dawn T Provenzale (dawn.provenzale@va.gov)

Version: 3 Date: 13 September 2012

Author’s response to reviews: see over
Dr. David Ozonoff, Editor

Environmental Health

Re: Manuscript 7964505457331856, A review of epidemiologic research assessments of United States Gulf War I era veterans

Dear Dr. Ozonoff,

September 13, 2012

Thank you for allowing us the opportunity to submit a revised version of the above-referenced manuscript. We are grateful for your kind assistance. We have carefully considered the suggestions of the reviewers, and have made the formatting changes that you requested. We feel that the scientific peer review process has substantially strengthened this manuscript, and hope that it is now acceptable for publication in Environmental Health.

A point-by-point response to the reviewers’ suggestions is included below.

With warm regards,

Rebecca McNeil

RESPONSE TO REVIEWERS

Reviewer 1

1. Inclusion and Exclusion Criteria Required. The authors need specific, a priori criteria to include or exclude specific studies in their review. The use of a ‘survey instrument’ should not be the only inclusion criteria. For example, include only cohort studies with representative study populations (demographic and military characteristics such as, active duty, National Guard / Reserve by branch of Service, rank, etc.).

Authors’ response: Thank you for pointing out the need to clarify our inclusion and exclusion criteria. We have expanded our description of the criteria as suggested.

2. Epidemiologic Perspective. Although understanding each survey instrument and comparing survey instruments across multiple study designs is one critical component of a Review article, understanding other key epidemiologic components across studies is equally important. For example, the research questions, methods, sampling frame(s), as well as study and comparison populations to ensure representativeness. Most important of all is the ability to generalizability [sic] the results. Unfortunately, the authors only address some of these issues in the current manuscript.

Additionally, the authors should consider describing any selection or measurement bias that may be unique to certain studies. Understanding any potential epidemiologic bias is critical to comparing studies. In other words, it doesn’t matter how good the survey instrument is, if selection bias or measurement error (how the outcome(s) was/were measured) is present.
Authors’ response: We agree that these are important points to consider in reviews of study methods and results. In lieu of the previous supplementary section summarizing the studies, we have added a tabulation of study characteristics, including the design, survey administration methods, sample selection procedures, research foci, and survey testing procedures (when available). This should allow the readers to more easily distinguish between the studies according to these large-scale characteristics, and evaluate their relevance to the more general population of veterans.

3. Limitations of Self-Reported Information. The authors should incorporate a discussion of the limitations of self-reported information, including demographic data, exposure(s) and health outcomes. As a rule, the highest quality epidemiologic data is objective and specific. This is true on the exposure side; as well as clearly defined, validated health events on the outcome side. Self-reported exposure(s) and self-reported health event(s) are neither.

Authors’ response: We have added a discussion of the challenges of self-reported data in the Validation section, which has been expanded at the suggestion of Reviewer 2 (comment 1). This is a topic of widespread concern, as reflected by several studies’ analyses of reliability and validity. We feel that this revised section, which contains a brief overview of the results of these analyses, more effectively conveys the community’s appropriate concerns regarding the quality of self-reported information.

4. Data Aggregation across Studies for Meta-analysis. Simply aggregating data across studies is inappropriate when individual studies are flawed, methods and study design are varied, or study populations do not represent or are not generalizable to a larger population of military members. I recommend deleting any references to meta-analytic methods.

Authors’ response: We agree that aggregating data across studies is fraught with problems. However, there is strong interest within the Gulf War community in combining existing Gulf War data sources, if possible, due in part to their relatively close temporal proximity to the war and arguably lower recall bias (see for example the VA’s Draft Gulf War Research Strategic Plan 2012-2016, http://www.va.gov/RAC-GWVI/docs/Minutes_and_Agendas/AppendixBDocument1.pdf). Thus, we would prefer to retain the existing references to meta-analysis. Because it is important to emphasize these issues of design flaws, heterogeneity, and generalizability, we have added a strongly worded caution regarding such, in hopes that the inevitable aggregation will be carried out with the appropriate careful considerations. We thank the reviewer for pointing out this excellent opportunity to stress the issues associated with this type of analysis.

5. The Millennium Cohort data presented in Tables One, Two, and Four is outdated and/or incomplete.

Authors’ response: Dr. Nancy Crum-Cianflone, Principal Investigator of the Millennium Cohort Study, has graciously provided us with current demographic data for use in Table 1, due to the absence of this information in the Millennium Cohort Study’s impressive body of literature. She also provided a review of the contents of Table 2 and 4. This updated information is reflected in the revised manuscript, and her kind support is noted in the Acknowledgements.

Reviewer 2
1. I found at times that the discussion was a little too broad. The section on Validation studies for example, situated just before the Conclusions, mentioned that the reliability and validity of seven studies ranged from fair to excellent agreement. Such terms are rather subjective, and lack detail. In my view it would be better to be more detailed, eg kappas (or correlation coefficients, or whatever the relevant measure) ranged from x (which is conventionally regarded by XXX, eg Landis and Koch, Biometrics 1977, 33, 159-174; in the case of kappa coefficients as being fair) to y (which is regarded by XXX as being moderate).

Authors’ response: As noted in the response to comment 3 from Reviewer 1, we have expanded the Validation section to include additional detail on the findings of reliability and validity analyses performed by the studies reviewed. We feel that the revised section more effectively conveys the desired information, and welcome the opportunity to add these details concerning important issues of data validity.

2. Similarly, it is certainly worthwhile to suggest that further studies use similar instruments, but what about when diverse instruments are used? For example, in the case of PTSD, 6 studies used the CAPS, 4 the PCL, given that there is not huge consensus, what should future studies use. Are there any empirical studies favouring one over the other, should studies go for the shortest or? On the subject of brief instruments, is there anything lost by using the SF-12 rather than the SF-36? A slightly more evaluative tone, looking at what should/could be used, as well as what has been used, might well aid planners of future/further studies. Obviously, this would only apply when a wide range of instruments has been used, when there is reasonable consensus as to measures of a particular concept, then those measures should arguably still be used.

Authors’ response: During the development of this manuscript, we considered taking a more evaluative and/or advisory tone, as you suggest. However, the number of domains that were studied using diverse instruments was large enough that extensive lengthening of the manuscript would have been required to comment on the relative merits of the instruments. We also felt that future investigators would benefit from familiarizing themselves with the candidate instruments that we listed, possibly in partnership with a psychometrician or survey researcher, in order to thoughtfully define the focus of their survey instruments. Thus, we have retained a more neutral descriptive tone.

We would like to note that the VA’s National Center for PTSD has compiled a list of PTSD screening questionnaires of potential interest to researchers. For each questionnaire included, this list includes a brief description, sample items, versions, validation studies, and references. It is located at http://www.ptsd.va.gov/professional/pages/assessments/list-screening-instruments.asp

Reviewer 3

1. Documentation of how this review is relevant to an environmental health journal.

Authors’ response: We have added emphasis that the health issues of Gulf War veterans have been of interest to researchers in diverse fields, including clinical medicine specialties, toxicology, occupational health, and epidemiology.
2. Review not only of measures used, but the psychometric properties of the surveys, especially the toxicant exposure variables.

Authors’ response: We agree that it is very important to consider the psychometric properties of surveys and survey instruments. A comprehensive review of the psychometric properties of all survey instruments referenced in this manuscript would be encyclopedic in its own right, and thus is not within the scope of our review. However, we have summarized the steps taken by each study to pilot test their surveys, for studies that published that information. That information is included in the new Table 1.

With regards to instruments to assess toxicant exposures: no study reviewed in this manuscript described the use of a validated instrument to assess this type of exposure. To the best of our knowledge, the Structured Neurotoxicant Assessment Checklist is the most-studied instrument available at this time; however, the results of validation studies for this instrument have not yet been published.

3. Review of findings within the neurotoxicant surveys would be more essential than a simple listing of surveys used in prior studies.

Authors’ response: We agree that a current comprehensive review of findings is needed. However, it is beyond the scope of this manuscript to include such. At this time, the most comprehensive summary of neurotoxicant exposure investigations may be found in the 2009 report of the Research Advisory Committee on Gulf War Veterans’ Illnesses. The Institute of Medicine’s Gulf War and Health series also contains some summary information on this topic.