

Justification A
Longitudinal Investigation of Gender, Health and Trauma (LIGHT) Survey
OMB FORM 2900-XXXX

A. JUSTIFICATION

1. Explain the circumstances that make the collection of information necessary. Identify legal or administrative requirements that necessitate the collection of information.

The National Center for PTSD (NCPTSD) was recently allocated funds by Congress to be used for research on prevention and treatment of PTSD. The original language of the legislation states the following: “The committee recognizes the importance of the VA National Center for PTSD in promoting better prevention, diagnoses, and treatment of PTSD. The Committee further recognizes the importance of this research for Veterans, their family members and those experiencing community violence. The Committee encourages the National Center to conduct further research on the effects of PTSD for veterans who live in communities affected by violence, particularly in low-income areas and communities of color.” In response to this, we have developed a study that aims at understanding the effects of trauma and community violence on US veterans, particularly women and racial minority veterans. The importance of this research cannot be over-stated as veterans living in high crime areas, racial minority veterans, and women veterans are often under-represented in studies querying mental health effects of trauma. Information gathered from this study will contribute to knowledge about factors that predict development of mental and physical health disorders as well as how individuals access and utilize healthcare. This information will directly inform intervention efforts aimed at prevention or treatment of chronic disorders such as PTSD, depression, and substance/alcohol use disorders, particularly in underserved portions of our veteran population. This study will further provide information as to what may interfere with Veterans’ ability to obtain needed health care for mental and physical health problems. This type of information can inform system-wide interventions that can maximize Veterans’ likelihood of receiving timely and evidence-based healthcare, preventing long-term health problems. As such, legal authority for this data collection is found under 38 USC, Part I, Chapter 5, Section 527, authorizing the collection of data that will allow for measurement and evaluation of the Department of Veterans Affairs Programs, the goal of which is improved healthcare for veterans.

2. Indicate how, by whom, and for what purposes the information is to be used; indicate actual use the agency has made of the information received from current collection.

The purpose of this study is to understand the cumulative effects of lifetime exposure to trauma and ongoing exposure to trauma such as community and intimate partner violence on Veterans’ mental and physical health, including its impact on the reproductive health of Veterans. To implement this research, VHA and entities working on behalf of VHA will conduct a nationwide longitudinal survey of Veterans residing in communities with varying levels of crime. Specifically, this longitudinal study will involve surveying Veterans regarding their life experiences, experiences within their neighborhood, mental health symptomatology, physical health, reproductive health, mental health service use, social support, and coping style four times over the course of approximately 1 year. We will contact a random sample of ~28,000 Veterans (~13,000 female and ~15,000 male) between the ages of 18 and 50 obtained from VA DoD Identity Repository (VADIR) to invite them to participate in this study, with the ultimate goal of achieving a baseline sample of ~4,000 Veterans. Given our primary aim to examine the role of community violence on outcomes, we will oversample for residency in high crime communities using zip codes to ensure that individuals living in these areas are invited to participate and are, therefore, represented in the study sample. We will also oversample rural communities using zip codes. Finally, as we are explicitly interested in under-represented populations in the larger Veteran population, we will also oversample racial minorities. Our response rate target for the survey is ~20%, which is consistent with other recent surveys of the Veteran population. After adjusting for potentially unusable or ineligible records (estimated at ~25%), we predict ~4,000 will complete the study.

Specific aims are:

Aim 1. To identify distinct health trajectories based on mental health, physical health, reproductive health, and functioning as a function of ongoing exposure to community violence.

Aim 2. To examine differences across gender and racial/ethnic groups in health trajectories as a function of current exposure to community violence.

Aim 3. To identify risk and protective factors that individually and interactively predicts health trajectories.

De-identified data will be made available to approved VA researchers. Researcher from other Federal Agencies and academic partners may also request access to de-identified data.

3. Describe whether, and to what extent, the collection of information involves the use of automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g. permitting electronic submission of responses, and the basis for the decision for adopting this means of collection. Also describe any consideration of using information technology to reduce burden.

We considered three methodologies for surveying participants: mail survey, phone survey and web-based survey. After comparing the relative advantages and disadvantages of these methods in terms of participants' convenience, privacy, data validity, logistics and cost, we chose a mail survey approach for this project. Accordingly, improved information technology will not decrease the burden on the public.

A mail survey has several advantages over a phone survey for this type of project. Unlike phone surveys, mail surveys can be completed at participants' leisure and so are more convenient. They also may offer greater privacy than phone interviews, which in some cases can be overheard. In addition, mail surveys reduce the potential for interviewer bias and may increase participants' comfort disclosing personal or sensitive information by providing an additional level of perceived anonymity. By following rigorous and recommended procedures (e.g., verified addresses, engaging packaging, repeat mailings), mail surveys can reach a greater proportion of the target audience than phone surveys, which encounter barriers related to unlisted numbers, decreased landline usage, answering machines, caller ID, and quick hang-ups. The main advantages to phone surveys, including lower demands for literacy and automated skip patterns, can be offset by using scales that have been validated with similar populations, carefully formatted instructions about which sections to complete and which to leave blank, and pre-testing the instrument before widespread use – all strategies we are using for the current project.

A mail survey also has several advantages over a web-based survey. First, mail surveys offer increased convenience for members of the target population who do not have easy access to the Internet. Further, it is easier to track who responds to a mail survey in order to ensure that our final sample is balanced on factors like gender. Whereas mail surveys can print an identification number or barcode on the survey, web-based surveys rely on participants accurately keying in an identification number. Web-based surveys may also raise Internet security concerns for some participants, reducing potential response rates. The main advantage of web-based surveys – automated skip patterns – can be addressed by using a well-designed mail survey, using the strategies described above.

4. Describe efforts to identify duplication. Show specifically why any similar information already available cannot be used or modified for use for the purposes described in Item 2 above.

To date, there have been no large-scale nationwide investigations of veterans that focus on veterans living in high crime areas. Additionally, this study will include a large sample of women and racial minority veterans to ensure their representation in this study as well as to answer novel questions related to reproductive health - an area that is largely understudied among Veterans.

5. If the collection of information impacts small businesses or other small entities, describe any methods used to minimize burden.

The collection of information involves randomly selected individuals in their residences, not small businesses or other small entities.

6. Describe the consequences to Federal program or policy activities if the collection is not conducted or is conducted less frequently as well as any technical or legal obstacles to reducing burden.

This study is critical to answer the question how cumulative exposure to trauma across the lifespan and current exposure to community violence and ongoing trauma impacts Veterans' physical and mental health, including reproductive health. The scientific rigor of the research program drives the frequency and type of data collected in order to most accurately answer the study questions, so that they may impact Veteran healthcare. Additionally, this study was specifically funded by Congress, which stipulates that funds need to be obligated (i.e., used or under contract to be used with a vendor) by September 2017. If the data collection is not done, not only will the scientific integrity and value of the research would be significantly diminished, we will not be able to use the funds as intended and allocated by Congress.

7. Explain any special circumstances that would cause an information collection to be conducted more often than quarterly or require respondents to prepare written responses to a collection of information in fewer than 30 days after receipt of it; submit more than an original and two copies of any document; retain records, other than health, medical, government contract, grant-in-aid, or tax records for more than three years; in connection with a statistical survey that is not designed to produce valid and reliable results that can be generalized to the universe of study and require the use of a statistical data classification that has not been reviewed and approved by OMB.

There are no such special circumstances.

8. a. If applicable, provide a copy and identify the date and page number of publication in the Federal Register of the sponsor's notice, required by 5 CFR 1320.8(d), soliciting comments on the information collection prior to submission to OMB. Summarize public comments received in response to that notice and describe actions taken by the sponsor in responses to these comments. Specifically address comments received on cost and hour burden.

The notice of Proposed Information Collection Activity was published in the Federal Register on August 8, 2017 (Volume 82, Number 37169, Page 37169). We received no comments in response to this notice.

b. Describe efforts to consult with persons outside the agency to obtain their views on the availability of data, frequency of collection, clarity of instructions and recordkeeping, disclosure or reporting format, and on the data elements to be recorded, disclosed or reported. Explain any circumstances which preclude consultation every three years with representatives of those from whom information is to be obtained.

Outside consultation is conducted with the public through the 60- and 30-day Federal Register notices. The study team includes a group of experts in clinical psychology, epidemiology, longitudinal survey design and analysis, women veterans, mental health, reproductive health, trauma and community violence. In addition, the study team has consulted with experts in these topics (e.g., reproductive health, healthcare utilization) outside the agency during the study and survey design process.

9. Explain any decision to provide any payment or gift to respondents, other than remuneration of contractors or grantees.

Each participant will be sent a \$5 incentive with the survey mailing, with the primary goal of increasing the response rate. Studies have shown that sending an up-front incentive increases the likelihood of response by 61%

and that sending a monetary incentive doubles the likelihood of response (Edwards, P., Roberts, I., Clarke, M., DiGuseppi, C., Pratap, S., Wentz, R., Kwan, I. & Cooper, R. [2007] Methods to increase response rates to postal questionnaires. *Cochrane Database of Systematic Reviews*, 4). While there is no consensus (or research to our knowledge) specifying the exact optimal amount, we find guidance in industry and marketing literature suggesting that “A \$5 incentive.... Boosts online survey response rates) <https://www.marketingcharts.com/industries/market-research-81847>. Further, \$5 incentives over smaller incentives is also supported by recent discussion generated on Focus Vision: “Researchers testing cash incentives in one dollar increments, ranging from 0 – \$10 discovered the following: The more cash they offered, the better the response rate to their survey. However, there was little difference between any of the \$0 – \$4 cash incentive conditions. Not until they offered \$5-\$8, was there a significant improvement over the lesser amounts. Interestingly, little difference was noted between any of the \$5-\$8 amounts. The \$10 condition showed the highest response rate at 26%.” <https://www.focusvision.com/blog/how-to-use-incentives-to-improve-response-rates-in-online-surveys/>. The NCBI of NIH suggests that \$10 is optimal. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5406995/> Given all of this research, we conclude that a \$5 pre-incentive is optimal given the added burden of this longitudinal study.

Participants also will receive an additional \$20 if they complete the survey. According to OMB’s incentives policy, incentives are most appropriately used in surveys with hard-to-find populations or respondents whose failure to participate would jeopardize the quality of the survey data, or in studies that impose exceptional burden on respondents, such as those asking highly sensitive questions. Given that our survey is specifically targeting hard to reach populations and includes very sensitive questions, a post survey incentive is appropriate.

This method will be the same at each of the four time points.

10. Describe any assurance of privacy, to the extent permitted by law, provided to respondents and the basis for the assurance in statute, regulation, or agency policy.

Information supplied to respondents in the form of a pre-survey letter and informed consent fact sheet will be reviewed and approved by the VA Boston Healthcare System’s Research and Development and Human Studies Committees (Institutional Review Boards) to insure the protection of study participants.

Information on these forms will become part of a system of records which complies with the Privacy Act of 1974. This system is identified as "Veteran, Patient, Employee and Volunteer Research and Development Project Records-VA (34VA11)" as set forth in the Compilation of Privacy Act Issuances via online GPO access at <http://www.gpoaccess.gov/privacyact/index.html>

11. Provide additional justification for any questions of a sensitive nature (Information that, with a reasonable degree of medical certainty, is likely to have a serious adverse effect on an individual's mental or physical health if revealed to him or her), such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private; include specific uses to be made of the information, the explanation to be given to persons from whom the information is requested, and any steps to be taken to obtain their consent.

As addressed above in question A.2, one of the primary aims of this investigation is to examine the effects of exposure to community violence and trauma on Veterans’ physical and mental health, including its impact on the reproductive health of Veterans. Accordingly, respondents will be questioned about experiences of trauma (both within and outside of the military), mental health symptoms and physical health symptoms. The investigative team has significant experience in the successful collection of information from victims of trauma, has a reputation for collecting such information in a compassionate manner and is sensitive to the possibility that participants may become distressed when recalling experiences of trauma. As addressed above in question A.10, all information provided to respondents, including the survey instrument, will be reviewed and approved by the VA Boston Healthcare System’s Research and Development and Human Studies Committees (Institutional Review Boards) prior to the initiation of data collection to insure the protection of study participants. The cover letter sent to participants with the survey instrument will explain the purpose of the survey and state that all analyses will be performed on aggregate-level data; no individual-level data will be reported. Additionally, participants will be

provided several contact numbers in the event they have questions about this study or if they experience any discomfort while participating in the study. Specifically, participants will be provided a phone number for the study Helpdesk, which can answer questions or put the participant in touch with the study investigators. They will also be provided with a Veteran resource list. To provide further privacy protection for participants, the investigative team will obtain a Certificate of Confidentiality from the National Institutes of Health (NIH) to protect identifiable research information from forced disclosure. This Certificate of Confidentiality will allow the investigative team to refuse to disclose identifying information on research participants in any civil, criminal, administrative, legislative, or other proceeding, thereby assuring confidentiality and privacy to participants.

12. Estimate of the hour burden of the collection of information:

- a. The number of respondents, frequency of responses, annual hour burden, and explanation for each form is reported as follows:**

Every effort has been made to minimize the data collection burden. The survey instrument was designed specifically to assess the critical constructs using the smallest number of reliable and valid items. Further, the survey instrument includes “skip-outs” so that respondents will not be required to respond to irrelevant questions. To this end, males and females will receive different survey versions so that they only receive questions that are relevant to them. Data collection will involve a mailed survey to randomly selected respondents, with a final target sample size of 4000. Pilot testing indicates that the survey instrument requires approximately 45 minutes to complete if a respondent answered all items (no skip-outs). Using these values, we compute the estimated burden as 16,000 hours, broken down as follows:

VA Form 10-XXXXX	No. of respondents	x No. of responses	x No. of minutes	÷ by 60 =	Number of Hours
Time 1 Survey	4,000	1	45		3,000
Time 2 Survey	4,000	1	45		3,000
Time 3 Survey	4,000	1	45		3,000
Time 4 Survey	4,000	1	45		3,000

- b. If this request for approval covers more than one form, provide separate hour burden estimates for each form and aggregate the hour burdens in Item 13 of OMB 83-I.**

See chart in subparagraph 12a above.

- c. Provide estimates of annual cost to respondents for the hour burdens for collections of information. The cost of contracting out or paying outside parties for information collection activities should not be included here. Instead, this cost should be included in Item 14.**

VA cannot make assumptions about the population of respondents because of the variability of factors, such as the educational background and wage potential of respondents. Therefore, VA/VHA uses general wage data to estimate the respondents’ costs associated with completing the information collection.

In accordance with the Bureau of Labor Statistics (BLS) May 2018 Occupational Wage Code Median Hourly, the mean hourly wage is \$24.98 based on the BLS wage code – “00-0000 All Occupations.” This information was taken from the following website: https://www.bls.gov/oes/current/oes_nat.htm

VA estimates the total annualized cost to respondents to be \$399,680 (16,000 burden hours x \$24.98 per hour).

Legally, respondents may not pay a person or business for assistance in completing the information collection. Therefore, there are no expected overhead costs for completing the information collection

13. Provide an estimate of the total annual cost burden to respondents or recordkeepers resulting from the collection of information. (Do not include the cost of any hour burden shown in Items 12 and 14).

- a. There are no capital, start-up, operation or maintenance costs.
- b. Cost estimates are not expected to vary widely. The only cost is that for the time of the respondent.
- c. There is no anticipated recordkeeping burden beyond that which is considered usual and customary.

14. Provide estimates of annual cost to the Federal Government. Also, provide a description of the method used to estimate cost, which should include quantification of hours, operation expenses (such as equipment, overhead, printing, and support staff), and any other expense that would not have been incurred without this collection of information. Agencies also may aggregate cost estimates from Items 12, 13, and 14 in a single table.

All costs for this data collection are included in funds directed by Congress to the NCPTSD for appropriation to this survey (\$1,019,719). These costs will cover the expense of 4 assessment intervals initially recruiting 28,000 participants with the estimated goal of 4,000 participants per assessment interval (4 intervals total) over the course of 1 year. There are no additional costs to the government for this activity.

15. Explain the reason for any burden hour changes or adjustments reported in items 13 or 14 of the OMB form 83-1.

This is a new collection and all burden hours are considered a program increase.

16. For collections of information whose results will be published, outline plans for tabulation and publication. Address any complex analytical techniques that will be used. Provide the time schedule for the entire project, including beginning and ending dates of the collection of information, completion of report, publication dates, and other actions.

Data will be tabulated and analyzed as follows. First, the data will be cleaned and checked for data entry errors, unusual variable distributions, potential errors due to skip patterns, the extent and patterns of missing data and outliers. For the items that are accompanied by multipoint Likert-type response formats ("strongly disagree"-to-"strongly agree"), the data will be checked by examining frequency distributions and calculating descriptive statistics. For dichotomous items, (e.g., "yes"/"no" responses), the probabilities of endorsement, or the proportion of respondents providing an affirmative response, will be reviewed. Next, we will construct scales using published scoring rubrics (when available) and compute Cronbach's alpha for internal consistency reliability, where appropriate.

After these initial steps in data checking, cleaning, and reduction, we will begin our analyses.

Aim 1. To identify distinct health trajectories based on mental health, physical health, reproductive health, and functioning as a function of ongoing exposure to community violence.

Hierarchical linear modeling (HLM; also called growth curve modeling, or multilevel or random coefficients regression) will address Aim 1. While traditional analytic methods conceptualize the dependent variable as an individual's state at a given time point, HLM treats the outcome as dynamic trends or trajectories providing for a deeper understanding of change over time. It also allows for identification of factors implicated in change over time vs initial status. Consistent with APA Task Force recommendations, we will attend to both statistical significance and effect sizes. Both sample design and nonresponse bias weights will help ensure results are generalizable to the

population. A criterion alpha of .05 will be used throughout. To address missing data, we use a full information maximum likelihood (FIML) estimation procedure to achieve reduced standard errors and more precise parameter estimates .

The Level-1 (within-subjects) component of the model will evaluate change over time in outcomes (physical health, mental health, reproductive health, functioning; separate models) producing estimates of initial status and change over time. Trauma and violence exposure will be included as a Level-2 predictor of the Level-1 change parameters. First, we will evaluate alternative unconditional change Level-1 outcome models (i.e., change in mental health symptoms without consideration of trauma/violence exposure) to determine the most reliable and powerful way to model time. Then, we will evaluate Level-2 predictor variables (i.e., trauma/violence exposure) of these Level-1 outcome change parameters.

Aim 2. To examine differences across gender and racial/ethnic groups in health trajectories as a function of current exposure to community violence.

Multilevel growth curve analyses similar to those described in Aim 1 will be conducted to evaluate the interactive effects of trauma/violence exposure and demographic factors (e.g., race/ethnicity, gender) on outcomes over time. Mean centered trauma/violence exposure and demographic scores and an interaction (product) term will be added as Level-2 predictors of the Level-1 change parameters. The regression coefficient for the interaction term predicting change over time will be examined to test the hypothesized interaction (i.e., moderation) predicting change in the outcome. Significant interactions will be explored graphically and statistically post-hoc in order to characterize the nature of the interaction.

Aim 3. To identify risk and protective factors that individually and interactively predict health trajectories.

Machine learning analyses (e.g., random forest) will be used to develop data-driven predictive models for outcomes of interest while incorporating the wealth of risk (e.g., ongoing violence exposure) and protective (e.g., social support) factor data collected at each study time point. The primary benefit of machine learning analysis compared with traditional regression is that it allows for the discovery of unknown pathways to outcomes of interest, as well as the identification of novel interactions between predictors, since it does not require a priori specification of predictors by the investigator. In the proposed study, machine learning can be used to develop predictive models of single time point outcomes (e.g., historical/current predictors at time 1 predicting suicidality at time 1), as well as outcome trajectories over the multiple time points of study data collection (e.g., historical/current predictors at time 1 predicting PTSD trajectory over the 4 time points of data collection). As predictors are added during new waves of data collection, changes in predictors can be examined, as well as the predictive ability of immediate versus longer-term predictors.

Because participants will be invited in two separate cohorts of ~14,000 each, the project's timeline calls for data collection to occur between September 2018 and March 2020. A complete final report will be submitted to the funding agency (approximately June 2020). Additionally, VA intends to publish this data in aggregate form. Dissemination of the study findings will include traditional academic mechanisms (e.g., articles published in peer-reviewed journals).

17. If seeking approval to omit the expiration date for OMB approval of the information collection, explain the reasons that display would be inappropriate.

The survey will include the expiration date.

18. Explain each exception to the certification statement identified in Item 19, "Certification for Paperwork Reduction Act Submissions," of OMB 83-I.

There are no exceptions.