SUPPORTING STATEMENT FOR:
VA PREPAREDNESS COMMUNICATIONS SURVEY
VA FORM 10-21086(NR)
2900-XXXX

B. Collections of Information Employing Statistical Methods

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

The sample for the proposed study will be pulled from the VA's VISTA system (Veterans Health Information and Technology Architecture). The VISTA system is a health care encounter-based record system; any Veteran who has a medical appointment at a VA site of care generates a record in the VISTA system complete with patient contact information, the date, time, location, and nature of the visit, the provider's name etc. Only the minimum necessary information from each encounter record will be used to field the proposed survey.

The sample will be ~~a nested design~~ stratified by respondent age (< 45, 45-64, and 64+) and respondent ZIP code including coastal/non-coastal regions as well as rural/urban regions.

The respondent universe for this study is comprised of two distinct populations: Veterans patients who are non-homeless, and Veterans patients who are homeless or transitionally housed (according to data available in the VISTA system).

Non-Homeless Veterans: The respondent universe includes Veterans who, at the time of fielding, both currently live in Veterans Integrated Service Networks (VISNs) 3, 4, or 5 and have ~~gotten~~ received care from a VA hospital or clinic in that same area at least once in the past 24 months. The three VISNs will be treated as one geographic location. At the time of sample design, VA appointment file data (obtained through VISTA) for this location identified 645,356 VA patients in a 12-month period.

VISNs 3, 4, & 5	Age under 45		Age 45 to 64		Age over 64		Total
	Rural	Urbanized	Rural	Urbanized	Rural	Urbanized	
Coastal	8,117	47,388	18,210	100,600	22,947	134,618	331,880
Non-Coastal	23,582	27,243	55,938	53,135	81,798	71,780	313,476
							645,356

Homeless or Transitionally Housed Veterans: A substantially smaller proportion of the respondent universe includes Veterans who, at the time of fielding, currently live in VISNs 3, 4, or 5 and have received care from a VA hospital or clinic in that same area at least once in the past 3 months and have received care from VA recorded as care delivered to the homeless. These Veteran patients have been identified through VISTA encounter data with the variable "primarystopcode" where a value of 528 = "Telephone/Homeless Mentally Ill" and 529 = "Health Care for Homeless Veterans (HCHV)". At the time of sample design, we identified 21,317 Veterans with these criteria over a 12-month time-frame, with roughly 5,000 Veterans meeting these criteria represented in any given 3-month period.

	Age under 45		Age 45 to 64		Age over 64		Total
	Rural	Urbanized	Rural	Urbanized	Rural	Urbanized	
VISNs 3, 4, AND 5	1,392	5,184	2,154	12,229	358	1,632	21,317

Coastal/non-coastal ZIP codes were identified using a database of coastal shoreline counties from the National Oceanic and Atmospheric Administration (NOAA) (http://coastalsocioeconomics.noaa.gov/coast_defined.html). Counties were matched to ZIP codes based on a database of ZIP codes from the U.S. Census Bureau. Urban and rural ZIP codes were defined using a database of ZIP codes from the U.S. Census Bureau.

There is not a sufficient number of transitionally housed or homeless Veterans to sample by strata, or to be included in the sample through random selection. Therefore, they will be excluded from the sampling process for the ~~nested~~ stratified sample design and will be sampled separately.

The sample size needed to obtain the target level of responses is dependent on estimated response rates. A wide range of reported response rates can be found based on the study population, survey mode, and use of financial incentives. One study of Veterans contacted by mail and phone reported response rates of 45% and 69% respectively (Duncan et al. 2005), while another mixed-mode survey of Veterans, very similar to our proposed use of mail, web, and phone components, reported an overall response rate of 34.3% (Eber et al. 2013). Another study conducted with initial monetary incentives found that $2, the amount we propose to provide initially, yielded a response rate of 47.7% (Griffin et al. 2011). Ultimately no published study directly mirrors ours and we will need to estimate response rates. Additionally, our response rates will be based not only on the willingness of the respondent, but also by the quality of the contact information.

Based on our fielding methodology which will use multiple contacts, modes, and cash incentives, and with good contact information in the sampling frame, we might expect response rates at or above 50%; however, due to the unknown quality of the sample data and the most recent study in 2013 by Eber achieving only at 34.3% response rate, we have estimated our response rate at 36% overall. Estimated response rates do vary by age (Wright et al., 2006). We did not find any literature on survey response rate differences by urban vs. rural or coastal vs. non-coastal. Therefore, our estimated response rate by cell is the following:

VISNs 3, 4, & 5	Age under 45		Age 45 to 64		Age over 64	
	Rural	Urbanized	Rural	Urbanized	Rural	Urbanized
Coastal	30%	30%	35%	35%	45%	45%
Non-Coastal	30%	30%	35%	35%	45%	45%

Our review of the literature found that only two studies (Sumner et al, 2001 and Elliot et al 2006) used a multi-stage probability sampling approach to contact homeless individuals. To our knowledge, all other surveys of the homeless population have been in-person interviews obtained through a convenience sample. Using the VISTA record of care delivered to a homeless Veteran presents a unique opportunity to obtain a probability sample of homeless or transitional Veterans. Since few studies like this have been done before, our estimated response rate is a conservative 5%.

2. Describe the procedures for the collection of information including: statistical methodology for stratification and sample selection, estimation procedure, degree of accuracy needed for the purpose described in the justification, unusual problems requiring specialized sampling procedures, and any use of periodic (less frequent than annual) data collection cycles to reduce burden.

To ensure we have enough survey responses to make statistical comparisons of response data, but not so many as to place undue burden on the study population, we have calculated our sample sizes based on several criteria and assumptions:

1. For all statistical hypotheses tests performed, we would like to have 95% confidence that we will not reject the null hypothesis when it is true. We have chosen a two-tailed test so that we can compare proportions no matter the direction of change ($\alpha = 0.05$).

2. We would like to have the statistical power to accept the null hypothesis when it is true 80% of the time ($\beta = 0.20$).

3. At its most conservative, the chance that a particular group of survey respondents will choose one survey response over another for a particular question is 50/50. Since no survey like this has been done before, our estimates will also be conservative ($p1 = 0.50$).

4. For comparing proportions of survey item responses between any two groups we would like to be able to determine significant differences of at least 6% ($p2 = 0.56$).

Our two-tailed sample size equation is the following:

$$n = \{ \frac{z_{1-\alpha/2}\sqrt{2\overline{pq}} + z_{1-\beta}\sqrt{p_1 q_1 + p_2 q_2}}{p_2 - p_1} \}^2$$

Where
$\alpha = 0.05$
$\beta = 0.2$

$Z_{1-\alpha/2} = 1.96$
$Z_{1-\beta} = 0.84$
$p_1 = 0.50$
$q_1 = 1-p_1 = 0.50$
$p_2 = 0.56$
$q_2 = 1-p_2 = 0.44$
$\bar{p} = ¿ (p_1 + p_2)/2 = (.50 + 0.56)/2 = 0.53$
$\bar{q} = ¿ 1- \bar{p} = ¿ 0.47$

Thus, to compare the differences in proportions between 2 mutually exclusive groups with our parameters we will need N=1,084 responses for each group, or 2,168 (2,172 with rounding).

VISNs 3, 4, AND 5	Age under 45		Age 45 to 64		Age over 64		Total
	Rural	Urbanized	Rural	Urbanized	Rural	Urbanized	
Coastal	181	181	181	181	181	181	1,086
Non-Coastal	181	181	181	181	181	181	1,086
Subtotal	362	362	362	362	362	362	
Total	724		724		724		2,172

As described above in section B.1, our overall estimated response rate is 36% with differing estimated response rates by strata.

VISNs 3, 4, AND 5	Age under 45		Age 45 to 64		Age over 64	
	Rural	Urbanized	Rural	Urbanized	Rural	Urbanized
Coastal	30%	30%	35%	35%	45%	45%
Non-Coastal	30%	30%	35%	35%	45%	45%

Dividing the desired number of completes per cell by the expected response rate, the level of sample we need to achieve 2,172 responses is 6,091. For example 181/0.3 = 603.

VISNs 3, 4, AND 5	Age under 45		Age 45 to 64		Age over 64		Total
	Rural	Urbanized	Rural	Urbanized	Rural	Urbanized	
Coastal	603	603	517	517	402	402	3045
Non-Coastal	603	603	517	517	402	402	3045

6091

With this number of completes we will be able to detect at least a 6% difference between survey responses calculated by rural/urban and coastal/non-coastal, and at least an 8.5% difference in proportions by the three age groups. The sample of 6,091 (rounded to 6,100) will be assembled into replicates, each representative of the sample as a whole, to contact no more than the minimum necessary to achieve 2,172 completed cases.

The 6,100 sample will be selected from the universe based on patient ZIP code, age, date of medical encounter, ZIP code of last medical encounter, and presence of a VA service-related disability rating. Active Duty Armed Service personnel are eligible to receive care at VA hospitals and clinics and may be in the sample if not purposefully excluded. The proposed study is a survey of Veterans only, and only Veterans have a listed service-related disability rating with the VA. By selecting our sample with that variable as a criteria, we are more likely to accurately select a Veteran patient who is eligible to complete this survey. With these criteria, sampling units with a medical encounter date within the past 24 months will be drawn randomly from the sample universe using the SAS v9.3 (Carey, NC) software program.

Using the homeless stop codes identified through VA encounter data presents a unique opportunity to obtain a probability sample of homeless or transitionally housed Veterans. Since no study like this has been done before, our estimated response rate is a conservative 5%. To get 100 completes, the goal for this study, we will need a sample size of 2,000 (100/0.05 = 2,000). The sample of 2,000 will be assembled into replicates, each representative of the sample as a whole, to contact no more than the minimum necessary to achieve 100 completed cases.

The 2,000 sample will be drawn from the 21,317 universe based on the presence of the variable "primarystopcode" where a value of 528 = "Telephone/Homeless Mentally Ill" and 529 = "Health Care for Homeless Veterans (HCHV)", date of medical encounter, ZIP code of last medical encounter, and patient ZIP code. With these criteria, sampling units with a medical encounter date within the past 3 months will be drawn randomly from the sample universe using the SAS program.

In total, our sample size is 8,100 (6,100 + 2,000), and our desired number of completes is 2,272 total (2,172 + 100).

This is a one-time pilot study. The study cannot be conducted less frequently.

3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

The proposed study was designed to maximize response rates. The method of using multiple modes of survey administration, in sequential order and ending with telephone CATI interviews of non-responders has been shown to improve response rates (Dillman et al., 2009). Additionally, the inclusion of a $2 cash incentive with the survey invitation has also been shown to improve response rates (Lesser et al., 2001). To address possible non-response to sensitive questions, we

have placed these at the end of the survey instrument. To address possible non-response due to respondent confusion, we have used simple sentence structure, small words where possible, avoided vernacular, and used validated questions from other public surveys where possible (e.g. U.S. Census Bureau). Lastly, non-response will be adjusted for in the weighting process for completed surveys. The study team will conduct a 4-stage weighting process:

1. *Base weights*. Base weights are the initial weights assigned to a given potential respondent in the sample. These weights are calculated as the inverse of the probability of selection for a given individual from within the population, by strata. These weights essentially represent the number of people that a given person within the sample initially represents. Given a random draw of individuals, the sample population is representative of the population as a whole once weights are applied with the base weights summing to strata and population totals.

2. *Propensity Score adjusted Non-response weights*. Although the base sample weight adjusts for varying probabilities of selection, all studies experience differential non-response across strata. To minimize potential bias in results, this differential response requires a post-field non-response weight to be calculated, to bring the final collected sample back to representing the original population. Altarum will use the generally accepted statistical practice of logistic regression to estimate propensity scores for respondents controlling for known factors among both the respondents and non-respondents. The propensity scores represent the probability of a given person to respond to the survey controlling for known socio-demographic characteristics. The inverse of the propensity scores will be multiplied by the corresponding base weights to bring the respondents in line with the total VISN 3, 4, and 5 Veteran populations.

3. *Post-Stratification Weights*. The application of propensity score adjusted non-response weights can lead to a mis-alignment of populations with some potentially excessive weights which skew the respondent populations data. Creation of a post-stratification weight will adjust the weights to ensure they best reflect the populations they represent.

4. *Final weights*. Final weights for each respondent will be calculated as the product of Base weight multiplied by the inverse of propensity score based non-response weight, multiplied by the post-stratification weight. Final weights will be used in conjunction with survey specific analytical techniques within SAS-callable SUDAAN which account for the complex survey design.

The survey team will calculate and report an adjusted response rate for the proposed study. The proposed study has three sequential modes of contact, ending with a Computer Assisted Telephone Interview (CATI) survey. Any study participant who does not respond to the first two contacts (web and mail) will be contacted for a CATI survey. Therefore, the study team will use adjusted response rate calculations for a CATI survey as suggested by the American Association of Public Opinion Research (AAPOR). Any survey completed in the first two contacts (web and mail) will be counted as a complete in that calculation.

4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents.

The survey will be administered in cycles of replicates. Collected cases will be monitored on an ongoing basis to identify any issues with the instrument, field protocols, or sample. The questionnaire instrument was tested with members of the study team's organization and reviewed by outside experts for edits. Additionally, where possible, we have used already validated questions from other public surveys (e.g. U.S. Census Bureau) in our questionnaire.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Aram Dobalian, PhD, JD
Principal Investigator
Director, Veterans Emergency Management Evaluation Center (VEMEC)
16111 Plummer St. MS-152
North Hills, CA 91343
818.891.7711 ext. 7182

Joe Swedorske
Lead Survey Sampling Methodologist
Altarum Institute
3520 Green Court, Suite 300
Ann Arbor, MI 48105
joe.swedorske#@altarum.org

Laura Nelson
Project Manager
Altarum Institute
2000 Duke St. STE 200
Alexandria, VA 22314
Laura.nelson@altarum.org

References

http://www.huffingtonpost.com/2012/11/29/superstorm-hurricane-sandy-deaths-2012_n_2209217.html

http://www2.va.gov/directory/guide/region.asp?isflash=0&ID=3http://www2.va.gov/directory/guide/region.asp?ID=3

Smith, Gavin. National Hazards Observer. Volume XXXVII Number 3 January 2013. http://www.colorado.edu/hazards/o/archives/2013/jan13_observerweb.pdf

Bochicchio, D. Memoradum to the Department of Veterans Affairs, Veterans Health Administration, Office of Emergency Management. 2014.

Claver M, Dobalian A, Fickel JJ, Ricci K, Horn-Mallers M. Comprehensive care for vulnerable elderly veterans during disasters. Archives of Gerontology and Geriatrics. 2013 Jan-Feb;56(1):205-13. 2012 Aug 16. [Epub ahead of print] PMID: 22901664.

Claver M, Friedman D, Dobalian A, Ricci K, Horn-Mallers M. The Role of Veterans Affairs in Emergency Management: A Systematic Literature Review. PLOS Currents: Disasters. 2012 Dec 12. Edition 1. PMC Journal – In process.

Dillman DA, Phelps G, Tortora R, Swift K, Kohrell J, Berck J, Messer BL. Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. Social Science Research 38 (2009) 1–18

Dobalian A, Callis R, Davey VJ. Evolution of the Veterans Health Administration's Role in Emergency Management Since September 11, 2001. Disaster Medicine and Public Health Preparedness. 2011 Sep;5 Suppl 2:S182-84. PMID: 21908696.

Dobalian A, Claver M, Fickel JJ. Hurricanes Katrina and Rita and the Department of Veterans Affairs: A Conceptual Model for Understanding the Evacuation of Nursing Homes. Gerontology. 2010;56(6):581-8. Epub 2010 Mar 24. PIMD: 20332609.

Dobalian A, Tsao JCI, Putzer GJ, Menendez SM. Improving rural community preparedness for the chronic health consequences of bioterrorism and other public health emergencies. Journal of Public Health Management and Practice. 2007. Sep/Oct;13(5) :476-80. PMID: 17762692.

Duncan P, Reker D, Kwon S, Lai SM, Studenski S, Perera S, Alfrey C, Marquez J. Measuring stroke impact with the stroke impact scale: telephone versus mail administration in veterans with stroke. Med Care. 2005 May;43(5):507-15

Eber S, Barth S, Kang H, Mahan C, Dursa E, Schneiderman A. The National Health Study for a New Generation of United States Veterans: methods for a large-scale study on the health of recent veterans. Mil Med. 2013 Sep;178(9):966-9

Elliott, M. N., Golinelli, D., Hambarsoomian, K., Perlman, J., & Wenzel, S. L. (2006). Sampling with field burden constraints: An application to sheltered homeless and low-income housed women. Field Methods, 18, 43–58.

Griffin JM, Simon AB, Hulbert E, Stevenson J, Grill JP, Noorbaloochi S, Partin MR. A comparison of small monetary incentives to convert survey non-respondents: a randomized control trial. BMC Med Res Methodol. 2011 May 26;11:81.

James, J.M. and Bolstein, R., "Large Monetary Incentives and the Effects on Mail Survey Response Rates," Public Opinion Quarterly 56:442-453 (1992).

Lesser VM, Dillman DA, Carlson J, Lorenz F, Mason RWillits F. Quantifying the Influence of Incentives on Mail Survey Response Rates and Their Effects on Nonresponse Error. Proceedings of the Annual Meeting of the American Statistical Association, August 5-9, 2001.

Putzer GJ, Koro-Ljungberg M, Duncan RP, Dobalian A. Preparedness of rural physicians for bioterrorist events in Florida. South Med J. 2013 Jan;106(1):21-6. doi: 10.1097/SMJ.0b013e31827caed2. PMID: 23263309.

Redline C, Oliver J, Fecso R. The Effect of Cover Letter Appeals and Visual Design on Response Rates in a Government Mail Survey. Conference Papers -- American Association For Public Opinion Research [serial online]. May 13, 2004;:N.PAG

Sumner GC, Andersen RM, Wenzel SL, Gelberg L. Weighting for Period Perspective in Samples of the Homeless. American Behavioral Scientist 2001; 45 (1): 80-104.

Tsao JCI, Dobalian A, Wiens BA, Gylys JA, Evans GD. Posttraumatic stress disorder in rural primary care: Improving care for mental health conditions following bioterrorism. Journal of Rural Health. 2006. Winter;22(1):78-82.

Wright SM, Craig T, Campbell S, Schaefer J, Humble C. Patient Satisfaction of Female and Male Users of Veterans Health Administration Services. J GEN INTERN MED 2006; 21:S26–32.

American Association of Public Opinion Research (AAPOR) Response Rate Calculator. Available at http://www.aapor.org/AAPORKentico/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx. Accessed on 4/15/2015.