Emergency Medicine Survey Sampling Methodology Report

Prepared by

Veteran Experience Office

Version 1 March 2020

Contents

VETERANS
EXPERIENCE
OFFICE

Executive Summary

The Emergency Medicine survey is designed to measure customer satisfaction after using either the Emergency Department or an Urgent Care department at a VHA facility. Any Veterans who recently obtained these services will be invited to participate in a brief online survey. The purpose of this report is to document the survey methodology and sampling plan of the Emergency Medicine survey.

Veterans experience data is collected by using an online transactional survey disseminated via an invitation email sent to randomly selected Veterans. The data collection occurs once a week within 7 days after Veterans have obtained emergency medicine services. The questionnaire is brief and contains general Likert-scale (a scale of 1-5 from Strongly Disagree to Strongly Agree) questions to assess customer satisfaction, as well as questions assessing the knowledge, speed, and manner of the interaction. After the survey has been distributed, recipients have two weeks to complete the survey and will receive a reminder email after one week.

The overall sample size for the Emergency Medicine population is determined so that the reliability of monthly survey estimate is at 3% Margin of Error at the 95% Confidence Level. Once data collection is completed, the participant responses in the online survey will be weighted so that the samples more closely represent the overall population. Weighting models will rely on the following: age, gender, facility complexity and urban/non-urban status.

This report describes the methodology used to conduct the Emergency Medicine survey. Information about quality assurance protocols, as well as limitations of the survey methodology, is also included in this report.

VETERANS
EXPERIENCE
OFFICE

Part I – Introduction

A. Background

The **Enterprise Measurement and Design** team (EMD) is part of the **Insights and Analytics** (I&A) division within the **Veterans Experience Office** (VEO). The EMD team is tasked with conducting transactional surveys of the Veteran population to measure their satisfaction with the Department of Veterans Affairs (VA) numerous benefit services. Thus, their mission is to empower Veterans by rapidly and discreetly collecting feedback on their interactions with such VA entities as NCA, VHA, and VBA. VEO surveys generally entail *probability* samples which only contact minimal numbers of Veterans necessary to obtain reliable estimates. This information is subsequently used by internal stakeholders to monitor, evaluate, and improve beneficiary processes. Veterans are always able to decline participation, and have the ability to opt out of future invitations. A *quarantine* protocol is maintained to limit the number of times a Veteran may be contacted, in order to prevent survey fatigue, across all VEO surveys.

Surveys issued by VEO are generally brief in nature, and present a low amount of burden to Veterans. A few targeted questions will utilize a human centered design (HCD) methodology, revolving aroundseven:

- Overall: (1) Satisfaction, (2) Confidence/Trust
- Service: (3) Effectiveness/Quality
- Process: (4) Ease/Simplicity, (5) Efficiency/Speed, (6) Equity/Transparency
- People: (7) Employee Helpfulness

Questions will focus on a specific aspect of a service process; spanning communication, applying for benefits, deliberation, and/or receipt of benefits. Structured questions directly address the pertinent issues regarding each surveyed line of business. The opportunity to volunteer open-ended text responses is provided within most surveys. This open text has been demonstrated to yield enormous information. Machine learning tools are used for text classification, ranking by sentiment scores, and screening for homelessness, depression, etc. Modern survey theory is used to create sample designs which are representative, statistically sound, and in accordance with OMB guidelines on federal surveys.

The Veteran Experience Office (VEO) has been commissioned by VHA to measure the satisfaction and experience of veterans that have recently obtained emergency medical services at a VHA facility. Sampled patients will be contacted through an invitation email. A link will be enclosed so the survey may be completed using an online interface, with customized patient information. The survey itself will consist of a handful of questions revolving around a human-centered design, focusing on such elements as trust, emotion, effective, and ease with the care they received.

B. Basic Definitions

Coverage	The percentage of the population of interest that is included in the sampling frame.
Measurement Error	The difference between the response coded and the true value of the characteristic being studied for a respondent.
Non-Response	Failure of some respondents in the sample to provide responses in the survey.
Transaction	A *transaction* refers to the specific time a Veteran interacts with the VA that impacts the Veteran's journey and their perception of VA's effectiveness in caring for Veterans.
Response Rate	The ratio of participating persons to the number of contacted persons. This is one of the basic indicators of survey quality.
Sample	In statistics, a data sample is a set of data collected and/or selected from a statistical population by a defined procedure.
Sampling Error	Error due to taking a particular sample instead of measuring every unit in the population.
Sampling Frame	A list of units in the population from which a sample may be selected.
Reliability	The consistency or dependability of a measure. Also referred to as *standard error*.

C. Application to Veterans Affairs

This measurement may bring insights and value to all stakeholders at VA. Front-line VA leaders can resolve individual feedback from Veterans and take steps to improve the customer experience; meanwhile VA executives can receive real-time updates on systematic trends that allow them to make changes.

1)   To collect continuous customer experience data from Veterans obtaining medical services from a VHA emergency department.
2)   To help field staff and the national office identify opportunities that may improve the experience of Veterans
3)   To better understand the reason that Veterans navigating the VHA emergency medicine department have a positive or negative experience.

Part II – Methodology

A. Target Population and Frame

The target population of the AMO surveys is all Veterans who have obtaing medical services at a VHA emergency medicine department in the seven days prior to sampling. Identification of these Veterans will occur via weekly data extracts from the VHA Corporate Data Warehouse (CDW), which houses the operational records of VHA.

B. Sample Size Determination

VETERANS
EXPERIENCE
OFFICE

For a given margin of error and confidence level, the sample size is calculated as below (Lohr, 1999). For population that is *large,* the equation below is used to yield a representative sample for proportions:

$$n_0 = \frac{Z_{\alpha/2}^2 \, pq}{e^2}$$

where

- $Z_{\alpha/2}$= 1.95, which is the critical Z score value under the normal distribution when using a 95% confidence level (α = 0.05).
- **p** = the estimated proportion of an attribute that is present in the population, with q=1-p.
  - o Note that pq attains its maximum when value p=0.5, and this is often used for a conservative sample size (i.e., large enough for any proportion).
- **e** = the level of precision achieved with the sample. Also referred to as the margin of error (MOE).

For a population that is relatively *small,* the finite population correction is used to yield a representative sample for proportions:

$$n = \frac{n_0}{1 + \dfrac{n_0}{N}}$$

Where

- n_0 = Representative sample for proportions when the population is large.
- **N** = Population size.

The margin of error surrounding the baseline proportion is calculated as:

$$Margin\,of\,error = z_{\alpha/2}\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{p(1-p)}{n}}$$

Where
- $Z_{\alpha/2}$= 1.95, which is the critical Z score value under the normal distribution when using a 95% confidence level (α = 0.05).
- **N** = Population size.
- **n** = Representative sample.
- **p** = the estimated proportion of an attribute that is present in the population, with q=1-p.

Sample sizes will be targeted to achieve 950 respondents overall each week. This sample size will allow for sustained sampling of the population over time and prevent oversampling of the same veteran in consecutive weeks. Table 2 shows the breakout of the population eligible for sampling and the target and fielded sample sizes for by geographic region and facility complexity level. The MOEs were calculated for two different confidence intervals (CI) to determine the sample's impact on the level of precision within each stratum.

Table 2. Sample Targets by Geography and Facility Complexity

	Email Population		Sample Size			Margin of Error	
	Total	Eligible for Sampling	Target	Fielded	Sampling Fraction	95% CI	90% CI

VETERANS EXPERIENCE OFFICE

Total	**18,4 31**	**18,1 00**	**9 50**	**4,7 50**	26%	3.1%	2.6%
Geographic Location							
Rural	4,0 25	3,9 53	2 50	1,2 50	32%	6.0%	5.0%
Urban/Other	14,4 06	14,1 47	7 00	3,5 00	25%	3.6%	3.0%
Facility Complexity Level							
1a	8,3 54	8,2 04	3 50	1,7 50	21%	5.1%	4.3%
1b	4,1 45	4,0 70	2 10	1,0 50	26%	6.6%	5.5%
1c	3,6 67	3,6 01	2 00	1,0 00	28%	6.7%	5.6%
2	1,1 82	1,1 61	1 00	5 00	43%	9.3%	7.8%
3	1,0 83	1,0 64	90	4 50	42%	9.8%	8.3%

C. Stratification

Stratification is used to ensure that the sample matches the population, to the extent possible, across sub-populations. The sample is stratified by geographic location of the veteran (rural/urban or other) and facility complexity level (1a, 1b, 1c, 2, or 3).   The sample will also use implicit sampling to assure that the sample reflects the demographic breakdown by veteran gender and age group (<30 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, 70+ years) within each explicit strata.

D. Data Collection Methods

The Emergency Medicine Surveys will be sample and conducted in tandem with the VEO outpatient survey. The sample frame will be drawn from the CDW every Tuesday. The initial survey invites will be sent the following day. After 7 days a reminder invite will be sent and the survey will close 14 days after the initial invitation.

Table 3. Survey Mode

Mode of Data Collection	Recruitment Method	Time After Transaction	Recruitment Period
Online Survey	Email Recruitment	Within 7 days after encounter	14 Days (Reminder after 7 Days)

VETERANS EXPERIENCE OFFICE

E. Reporting

Researchers will be able to use the VSignals platform for interactive reporting and data visualization. Trust, Ease, Effectiveness, and Emotion scores can be observed for each). The scores may viewed by Age Group, Gender, and Race/Ethnicity in various charts for different perspective. They are also depicted within time series plots to investigate trends. Finally, filter options are available to assess scores at varying time periods and within the context of other collected variable information.

Recruitment is continuous (weekly) but the results from several weeks may be combined into a *monthly* estimate for more precise estimates, which is the recommended reporting level. Weekly estimates may include minor distortions but allow analysts to review scores more quickly and within smaller time intervals. Weekly estimates are less reliable for small domains, and should only be considered for aggregated populations. Monthly estimates will have larger sample sizes, and therefore higher reliability set to a 3% MOE at the 95% Confidence level. All estimates are also weighted in real time on the platform for improved representation and less bias (non-response and coverage, see section G on Sample Weighting) but the weights can introduce distortions when looking at short time windows. Quarterly estimates are the most precise, but will take the greatest amount of time to obtain (12 weeks of collection). However, Quarterly estimates are the most suitable for the analysis of small populations (e.g. VAMC, Female Veterans 18-29, etc.).

F. Quality Control

To ensure the prevention of errors and inconsistencies in the data and the analysis, quality control procedures will be instituted in several steps of the survey process. Records will undergo a cleaning during the population file creation. The quality control steps are as follows.

1. Records will be reviewed for missing sampling and weighting variable data. When records with missing data are discovered, they will be either excluded from the population file or put into separate strata upon discussion with subject matter experts.
2. Any duplicate records will be removed from the population file to both maintain the probabilities of selection and prevent the double sampling of the same veteran.
3. Invalid emails will be removed.

The survey sample loading and administration processes will have quality control measures built into them.

1. The extracted sample will be reviewed for representativeness. A secondary review will be applied to the final respondent sample.
2. The survey load process will be rigorously tested prior to the induction of the C&P Survey to ensure that sampled customers is not inadvertently dropped or sent multiple emails.
3. The email delivery process is monitored to ensure that bounce-back records will not hold up the email delivery process.

The weighting and data management quality control checks are as follows:

1. The sum of the weighted respondents will be compared to the overall population count to confirm that the records are being properly weighted. When the sum does not match the population count, weighting classes will be collapsed to correct this issue.

VETERANS
EXPERIENCE
OFFICE

2. The unequal weighting effect will be used to identify potential issues in the weighting process. Large unequal weighting effects indicate a problem with the weighting classes, such as a record receiving a large weight to compensate for nonresponse or coverage bias.

G. Sample Weighting, Coverage Bias, and Non-Response Bias

A final respondent sample should closely resemble the true population, in terms of the demographic distributions (e.g. age groups). One problem that arises in the survey collection process is **nonresponse**, which is defined as failure of selected persons in the sample to provide responses. This occurs in various degrees to *all* surveys, but the resulting estimates can be distorted when some groups are actually more or less prone to complete the survey. In many applications, younger people are less likely to participate than older persons. Another problem is **under-coverage**, which is the event that certain groups of interest in the population are not even included in the sampling frame. They cannot participate because they cannot be contacted: those without an email address will be excluded from sample frame. These two phenomena may cause some groups to be over- or under-represented. In such cases, when the respondent population does not match the true population, conclusions drawn from the survey data may not be reliable, and are said to be **biased**.

Survey practitioners recommend the use of sampling weighting to improve inference on the population. This will be introduced into the survey process as a tool that helps the respondent sample more closely represent the overall population. Weighting adjustments are commonly applied in surveys to correct for nonresponse bias and coverage bias. As a business rule will be implemented to require callers to provide email address, the coverage bias for this survey is expected to decrease. In many surveys, however, differential response rates may be observed across age groups. In the event that some age groups are more represented in the final respondent sample, the weighting application will yield somewhat smaller weights for this age group. Conversely, age groups that are underrepresented will receive larger weights. This phenomenon is termed *non-response bias correction* for a single variable. Strictly speaking, we can never know how non-respondents would have really answered the question, but the aforementioned adjustment calibrates the sample to resemble the full population – from the perspective of demographics. This may result in a substantial correction in the resulting weighting survey estimates when compared to direct estimates in the presence of non-negligible sample error (non-response bias).

Because the email population will have different demographics than the overall population, the initial sample will be selected in a manner from the frame so that the final respondent sample resembles the overall population. Stratification may also adjust for non-response (occurring when certain subpopulations are less prone to participate). Targets will be established for every permutation of the following stratification variables. As such, population values will be collected and recorded by VEO for every data collection period.

The stratification scheme above will result in a representative sample with respect to to the full population. Weighting will then be applied so that the sample is more fully matched to the population. Sample weights will be generated for monthly estimates.

Since 85% of older Americans utilize email (Choi & Dinitto, 2013), we can presume that a large share of veterans chose not to share their email address with VHA or are simply unware of that option. It is assumed that the level of patient satisfaction is not directly related to their email status (Missing at Random). Since age and gender have been observed to be strong predictors of patient satisfaction in other VA health surveys, the stratification and weighting methodology outlined above will adequately compensate for any bias introduced by the incomplete frame of population.

VETERANS
EXPERIENCE
OFFICE

Weighting will utilize cell weights in real time. To make this possible, targets will be based on the previous month's population. With each query on the VSignals platform for each respondent by dividing the target for a cell by the number of respondents in the cell. The weighting scheme will include, where possible all the variables used for explicit stratification, However, cells will be collapsed if the proportion of the population is insufficient to reliably achieve a minimum of 3 completes per month. As a result, weights may be more comprehensive for larger population segments. For instance, in the VA, women are a smaller proportion of the populations. Therefore, woman will have more collapsed cells than men.

As part of the weighting validation process, the weights of persons in age and gender groups are summed and verified that they match the universe estimates (i.e., population totals). Additionally, we calculate the *unequal weighting effect*, or UWE (see Kish, 1992; Liu et al., 2002). This statistic is an indication of the amount of variation that may be expected due to the inclusion of weighting. The unequal weighting effect estimates the percent increase in the variance of the final estimate due to the presence of weights and is calculated as:

$$UWE = 1 + cv^2_{weights} = 1 + \left(\frac{s}{\overline{w}}\right)^2$$

where

- **cv** = coefficient of variation for all weights w_{ij}.
- **s** = sample standard deviation of weights.
- $\overline{w}$ = sample mean of weights, $\overline{w} = ¿ \dfrac{1}{n} \sum_{ij} w_{ij}$.

H. Quarantine Rules

VEO seeks to limit contact with Veterans as much as possible, and only as needed to achieve measurement goals. These rules are enacted to prevent excessive recruitment attempts upon Veterans. VEO also monitors veteran participation within other surveys, to ensure veterans do not experience survey fatigue. All VEO surveys offer options for respondents to opt out, and ensure they are no longer contacted for a specific survey. VEO also monitors Veteran participation within other surveys, to ensure Veterans do not experience survey *fatigue.* Finally, all VEO surveys offer options for respondents to opt out, and ensure they are no longer contacted for a specific survey.

Table 4. Quarantine Protocol

Quarantine Rule	Description	Elapsed Time
Repeated Sampling for AMO Survey	Number of days between receiving/completing online survey, prior to receiving email invitation for a separate C&P experience	30 Days
Other VEO Surveys	Number of days between receiving/completing online survey and becoming eligible for another VEO survey	30 Days
Prioritization	Prioritization is based on the observed sample sizes.	N/A

VETERANS
EXPERIENCE
OFFICE

Opt Outs	Persons indicating their wish to opt out of either phone or online survey will no longer be contacted.	N/A

Part III – Assumptions and Limitations

A. Coverage Bias

Since the VEO Emergency Medicine surveys are email only, there is a large population of patients that cannot be reached by the survey. Veterans that lack access to the internet or do not use email may have different levels of Trust and satisfaction with their service. However, the majority of Veterans that do not share their email addresses do so because they did not have an opportunity to provide the information, or they elected not to share their email address. As such, it is thought that Veterans in this latter category do not harbor any tangible differences to other Veterans who do share their information. In order to verify this, VEO plans to execute a coverage bias study to assess the amount of coverage bias due, and derive adjustment factors in the presence of non-negligible bias.

VETERANS
EXPERIENCE
OFFICE

Appendix 3. References

Choi, N.G. & Dinitto, D.M. (2013). Internet Use Among Older Adults: Association with Health Needs, Psychological Capital, and Social Capital. *Journal of Medical Internet Research*, 15(5), e97

Kalton, G., & Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics*, 19(2), 81-97.

Kish, L. (1992). Weighting for unequal P. *Journal of Official Statistics*, 8(2), 183-200.

Kolenikov, S. (2014). Calibrating Survey Data Using Iterative Proportional Fitting (Raking). *The Stata Journal,* 14(1): 22–59.

Lohr, S. (1999). *Sampling: Design and Analysis* (Ed.). Boston, MA: Cengage Learning.

Liu, J., Iannacchione, V., & Byron, M. (2002). Decomposing design effects for stratified sampling. *Proceedings of the American Statistical Association's Section on Survey Research Methods*.

Wong, D.W.S. (1992) The Reliability of Using the Iterative Proportional Fitting Procedure. *The Professional Geographer,* 44 (3), 1992, pp. 340-348

VETERANS
EXPERIENCE
OFFICE