Section B - Statistical Methods

National Blood Collection and Utilization Survey

OMB No. 0990-0313 - Revision

B. Collection of Information Employing Statistical Methods.

1. Respondent Universe and Sampling Methods

Respondent Universe. The population of inference for the 2021 NBCUS will be all blood collection and utilization facilities in the U.S. The target population for the 2021 NBCUS will consist of all blood collection centers, cord blood banks and all hospitals subject to certain ownership, service and location criteria. Some practical restrictions were also placed on the target population – specifically, hospitals reporting fewer than 100 inpatient surgeries per year were excluded since they contribute little to either collections or blood product utilization.

Sampling Frame and Sample Design. We used the American Hospital Association's (AHA) annual survey of hospitals, U.S. Food and Drug Administration's (FDA) Blood Establishment Registration database, and America's Blood Center's (ABC) contact list together to construct a sampling frame that covers virtually all collection, processing and transfusion of blood. Hospitals on the AHA file were included in the 2021 NBCUS sampling frame subject to ownership, services, location and surgical volume criteria as follows:

1) Ownership – The AHA ownership (CNTRL) must be:
   - Veterans Affairs (45) or
   - Other non-Federal
   - State (12)
   - County (13)
   - City (14)
   - City-county (15)
   - Hospital district or authority (16)
   - Church operated (21)
   - Other (23)
   - Individual (31)
   - Partnership (32)
   - Corporation (33)

   AND

2) Services – The AHA primary service (SERV) must be:
   - General medical and surgical (10) or
   - Surgical (13) or
   - Cancer (41) or
   - Heart (42) or
   - Obstetrics and Gynecology (44) or

Eye, ear, nose and throat (45) or
Orthopedic (47) or
Children's general (50) or
Children's orthopedic (57)

AND

3) Location – Located within the 50 United States (or the District of Columbia).

AND

4) Surgical Volume – The hospital must conduct 100 or more inpatient surgeries per year.

Hospitals on that could be matched to the AHA file (i.e., hospitals found on both files) were included in the NBCUS frame and sample subject to the information available in the AHA file and the eligibility criteria described above.

Following the previous 2005, 2007, 2009, 2011, 2013, 2015, 2017, and 2019 National Blood Collection and Utilization Surveys, we drew a stratified, single stage sample of blood banks and hospitals with equal probability within stratum. We stratified hospitals on the AHA file by size (annual inpatient surgical volume), and selected hospitals in the larger size strata with certainty. Table B1-1 below gives the total population, sample size and sampling rate for the various types of facilities.

Table B1 – 1 Total population, sample size and sampling rate by type of facility

Type of facility	Total population	Sample	Sampling rate (%)
Hospitals (AHA) – annual inpatient surgical volume			
100-999	1,672	669	40.00
1,000-1,3999	350	350	100.00
1,400-2,399	589	589	100.00
2,400-4,999	682	682	100.00
5,000-7,999	281	281	100.00
≥8,000	237	237	100.00
Community-based blood collection center	53	53	100.00
Hospital-based blood collection center	90	90	100.00
Total	3,954	2,951	74.6

* Institutions such as the American Red Cross will have their central data repository (ARCNET) reporting for all Red Cross centers. Therefore, the number of blood centers sampled does not correspond to the total number of blood centers in the United States.

The response rates for the 2019 NBCUS were 94.3% (50/53) for community-based blood collection facilities, 84.4% (76/90) for hospital-based blood collection facilities, and 76.2% (2140/2808) for transfusing hospitals. Based on the previous three iterations of the NBCUS, we expect an overall response rate of almost 80%

across all types of facilities.

As with all establishment samples, we anticipate that units on the sampling frame (whether they are sampled or not) can merge with one another, split into multiple units, etc. Such events have implications for calculating overall probabilities of selection. We plan to implement procedures that can be reflected in the table capture the information relevant to calculating correct overall probabilities of selection and that also could deal with the phenomenon of sampled units reporting for different organizational levels.

2. Procedures for the Collection of Information

Initial Contact. An introductory letter will be sent to the Director of Transfusion Services of each sampled institution. The letter describes the purpose of the survey, the authority for data collection, and provides a pre-notification on the types of information that will be requested on the questionnaire. This will give institutions the opportunity to gather information from 2021 to ease in completing the survey. The introductory letter will have a postage paid tear off section asking for confirmation of the name and contact information of the person who would most likely complete the survey at the specified institution. Verification of the appropriate contact within the selected hospital or blood bank will help increase response rates.

Survey Mailing. To maximize response rates as well as the quality of response data, a survey packet along with a prepaid envelope will be mailed. This packet will contain a unique, secure link to the online survey instrument. The person assigned to complete the questionnaire will be given a unique login and password. A cover letter co-signed by the Department of Health and Human Services (HHS) and Blood Organ and Other Tissue Safety (BOOTS) Office will accompany the survey packet. The cover letter will come from a significant HHS official such as the Assistant Secretary of Health or the Director of the BOOTS Office.

Follow-up. Two weeks after the initial mailing, follow-up will be made to non-respondents with a postcard. The postcard will provide information on the survey and ask again for participation. In addition, the recipients will be provided with a Survey Helpline information. The toll-free helpline will field inquiries related to the survey and will be available 24 hours a day. Two weeks after the second mailing a phone call will be made. The phone call will seek to determine obstacles to completing the survey and offer to help in any way possible. Two weeks following the phone call a letter and a copy of the survey instrument link will be sent with a request to complete

the questionnaire. Frequently, the first copy of the survey goes astray, and it is the subsequent mailing that will stimulate action.

Monitoring Data Collection and Quality Control. A survey receipt control system will be used to track and monitor distribution of questionnaires and responses, helping to ensure that actions are taken in a timely manner to maximize response rates. All of the sampled institutions will be entered into a Microsoft Access database to track the mailing, receipt and processing of the questionnaires. When a questionnaire is returned, it will be entered into a table that tracks its processing status. Key to obtaining good response rates and complete data in this type of study is developing a rapport with the individual(s) who will be completing the questionnaire and ensuring that the survey gets to the right person who has the knowledge to respond. The use of the tracking system will assist with this process. In addition, institutions that do not respond will be offered the opportunity to complete the abbreviated version of the survey. This will help obtain critical information from as many institutions as possible. The tracking system will be the vehicle for follow-up of participation status.

Coding schemes will be used to classify original data from source documents into codes that are machine-readable and appropriate for planned analysis. When an unusual response is recorded that does not appear to fit into the existing coding scheme, it will be reviewed by the data manager to determine how best to handle it. The problem will be passed along to the project director if the data manager is unable to resolve it. COED, a proprietary codebook, and machine-edit PC software will be used to create codebooks and define computerized checks. Manual editing will be performed simultaneously with coding. Manual editing of study documents includes checking for illegible answers, incorrectly followed skip instructions, items not answered, responses outside the ranges of acceptable answers specified in the codebook (range check), responses of an incorrect character length, inappropriate responses and inconsistency between answers. In addition, once initial batches of data have been keyed from study documents, a manual edit, or proofreading, of that data will be completed to ensure that all variables have been keyed into the correct record and column positions.

Verification of coding is an important tool that will be used for quality control. The verification sample of at least 10 percent or higher will be a check of the accuracy of coding and will be carried out by independent recoding. Once the discrepancy or

error rate falls to an acceptable level, the verification rate will be cut, but will never be lower than 10 percent.

Prepared and coded survey data will then be keyed and verified. Data entry and verification will be performed using customized PC-based commercial software. Data entry programs will be set up to contain checks for completeness of keyed data, including range checks for pre-coded items and skip pattern verification. After the raw data files have been edited and updated, they will be converted into SAS data files.

Weighting. Base weights will be calculated for each unit as the reciprocal of its overall probability of selection. These base weights will then be adjusted for non-response. We will use sampling strata as initial non-response adjustment cells, which can be further refined through the use of Chi-squared Automatic Interaction Detector (CHAID) or other response propensity modeling software to incorporate other variables from the sampling frame that appear related to response propensity. Minimum cell sizes and non-response adjustment factors will be considered in the final non-response cell definition, in order to avoid unnecessarily large increases in variance due to differential weighting. The use of post stratification raking or calibration to adjust the weights to one or more known or estimated population totals available from the sampling frame, the annual AHA survey data file and the member list will be considered. These adjustments have the effect of increasing the precision of estimates, while matching known population counts.

Imputation. All data items will be checked for internal consistency as part of the data cleaning process and as a prelude to imputation. Missing data will be imputed in continuous (i.e., interval or ratio-level) variables via regression or time series models that take into account previously reported data for the same unit (when available), as well as previously and presently reported data for similar, responding units. Separate imputation models will be used for blood centers and hospitals including volume-related variables (e.g., collection, transfusion and/or surgical counts) as predictor variables. Imputation of nominal or ordinal-level variables, to the extent warranted or desired will be achieved through the use of proprietary of AutoImpute software, which uses a combination of continuous and categorical data modeling, hot-deck imputation and successive iterations to impute such items, while acknowledging questionnaire skip patterns, etc. All imputed data items will be checked for internal consistency using the same routines as the data cleaning process and imputation

flags will be provided in the analytic dataset to distinguish imputed from reported values.

3. Methods to Maximize Response Rates and Deal with Nonresponse

The CDC's established relationships in the blood collection and transfusion community, combined with lessons learned from conducting the 2005–2019 NBCUSs, will help enhance participation in the 2021 NBCUS. Announcements will be made at Annual Meetings to notify the community of the upcoming survey. BOOTS will use its communication vehicles, distributed daily, weekly and monthly, to help recruit blood centers and member hospitals and to provide updates and information on the 2021 NBCUS.

A pre-notification letter asking for confirmation of appropriate contact person at each institution, a follow-up postcard at two-weeks after the initial survey mailing, a telephone call at four weeks, and a follow-up letter and survey mailing at six weeks, along with the toll-free helpline will help us achieve the desired response rate.

Despite the methods described above, we still expect some eligible sampled units to be non-respondents. We will deal with non-response and its potential impact on survey estimates through a combination of weight adjustments and non-response bias analysis. As described above (Section B2, Weighting), base weights will be adjusted for non-response using CHAID or other response propensity modeling software to incorporate variables from the sampling frame (other than the sampling strata) that appear related to response propensity.

A non-response bias analysis will then take advantage of the detailed information available for both responding and non-responding sampled units from the annual AHA survey data file and the member list to assess the potential for non-response bias due to both unit (i.e., complete) and item (i.e., item specific) non-response. We will use differences in unit and item response rates across the various detailed data items, both before and after weight adjustments, as a proxy for the potential for non-response bias. We will include a summary of the results of this non-response bias analysis in technical and analytic reports.

4. Tests of Procedures or Methods to be Undertaken

The proposed 2021 survey instrument and data collections procedures are by and large the same as the 2005, 2007, 2009 2011, 2013, 2015, 2017, and 2019 NBCUSs which received OMB approval and achieved satisfactory results. For this reason, we will not be conducting pilot tests. However, consultation was sought from

individuals within HHS, blood center staff familiar with center operations, and experts in transfusion medicine. Additionally, we have requested input from selected stakeholders on the survey instrument by requesting feedback on the new questions.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The statistician responsible for the survey sample design is:

Matt Sapiano, PhD
Statistician
Lantana Consulting Group (CTR)
Division of Healthcare Quality Promotion
Centers for Disease Control and Prevention
1600 Clifton Rd.
Atlanta, GA 30329
Tel: 404.718.5761
lof6@cdc.gov

Data collection, analysis, and quality control will be under the supervision of:

Rebecca Free, MD, MPH
Medical Epidemiologist
Division of Healthcare Quality Promotion
National Center for Emerging and Zoonotic Infectious Diseases
Centers for Disease Control and Prevention
1600 Clifton Rd, NE MS V18-4
Atlanta, GA 30329
ksz4@cdc.gov

Sridhar Basavaraju, MD, FACEP
Director - Office of Blood, Organ, and Other Tissue Safety
Division of Healthcare Quality Promotion
National Center for Emerging and Zoonotic Infectious Diseases
Centers for Disease Control and Prevention
1600 Clifton Road, NE MS V18-4
Atlanta, GA 30329
Tel: 404.498.0729
etu7@cdc.gov

James Berger, MS, MT (ASCP), SBB
Senior Advisor for Blood and Tissue Safety
Office of Infectious Diseases and HIV/AIDS Policy
Office of the Assistant Secretary for Health
U.S. Department of Health and Human Services
330 C Street, SW, Suite L127
Washington, DC 20024
Tel: 202.795.7608
james.berger@hhs.gov