**SUPPORTING STATEMENT**

**Part B**

**Pilot Test of the Proposed Hospital Survey on Patient Safety Culture**

**Version 2.0**

**Version: March 20, 2015**

Agency for Healthcare Research and Quality (AHRQ)

Table of Contents

[B. STATISTICAL METHODS 2](#_Toc378946828)

[1. Potential Respondent Universe and Sample Selection Method 2](#_Toc378946829)

[2. Information Collection Procedures 4](#_Toc378946830)

[3. Methods to Maximize Response Rate 4](#_Toc378946831)

[4. Tests of Procedures 5](#_Toc378946832)

[5. Statistical Consultation and Independent Review 5](#_Toc378946833)

# B. STATISTICAL METHODS

## 1. Potential Respondent Universe and Sample Selection Method

**Cognitive Interview Participants**

Cognitive interviews will be done on a) the draft Hospital Survey on Patient Safety Culture Version 2.0 (HSOPS 2.0) before larger-scale pilot testing is done and b) the supplemental item set on Health Information Technology (Health IT) Patient Safety. Cognitive interview participants will be selected from hospitals that will vary by type, size and geographic location. We will use participant recruitment methods similar to those used for the other patient safety culture surveys, such as sending research participation flyers to hospitals. We aim to conduct cognitive testing with a total of 72 hospital providers and staff. We will conduct cognitive interviews with 54 English speaking hospital providers and staff across the range of positions found in hospitals, from physicians and nurses to dietitians and housekeepers. We will also conduct an additional 18 cognitive interviews with Spanish speaking participants for testing the Spanish translation of the instrument. No special selection procedures will be used to select specific participants within hospitals with the exception of ensuring the appropriate distribution across staff positions and hospital type.

**Pilot Test and Bridge Study Sample**

A sampling frame of potential hospitals will be developed in consultation with AHRQ staff and representatives from the TEP (see Attachment H for a list of TEP members). These persons will provide points of contact in hospitals that might be willing to participate.

The purposes of the overall sample design are three-fold: (1) to obtain enough pilot test data at both the hospital site level and the individual respondent level to ensure sufficient sample size (*n*) for examining the multilevel psychometric properties of the data, (2) to include a variety of hospitals that differ in type, size, and geographic region, and (3) to maximize our power to detect differences between scores on the original and revised version of the *Hospital Survey on Patient Safety Culture*.

Since the goals are to examine the psychometric properties of the survey and to detect differences between survey versions, not to produce national estimates, purposive sampling will be used. Purposive sampling will ensure adequate variability on important hospital characteristics given the small number of hospitals included in the pilot test. It should be noted that the reason for including hospitals of different types is not to compare survey results across the types but rather to ensure that there is some representativeness of hospitals by type. The final survey will be publicly available for use by all types of hospitals.

As shown in Table 1, we propose that data be collected from 40 hospitals – ranging from four to ten hospitals within each of the six categories of hospitals by type and size. We do not propose to evenly distribute the number of hospitals across the six categories but rather propose distributing the hospitals into these categories based on the distribution among hospitals using the HSOPS and submitting to the HSOPS Comparative Database Report.

We propose a split ballot design in which half of respondents in each hospital receive the original HSOPS survey instrument and the other half receives the draft HSOPS 2.0 and supplemental item set. This will allow assessment of differences within as well as across hospitals. We will require a sample of approximately 500 providers and staff in each hospital for within-hospital comparison of the two surveys, 250 will receive the original HSOPS and 250 will receive the draft HSOPS 2.0 and supplemental item set. The data from the 250 providers and staff who receive the draft HSOPS 2.0 and supplemental item set will be used for the pilot study analyses to revise and finalize HSOPS 2.0 and supplemental item set. Because we require such a large sample within each hospital, we will target only hospitals with 49 or more beds. For hospitals with fewer than 500 providers and staff, we will conduct a census in the hospital (assuming on average 375 providers and staff in these hospitals). As shown in Table 2, this will yield a total of 18,375 sample members assuming all 40 hospitals participate. Assuming a response rate of 50 percent, this will yield a total of 9,188 completed questionnaires.

Participating hospitals will deliver the sampled provider and staff lists to AHRQ or its contractor. Hospital providers and staff will be stratified by work area and staff positions, and the 500 individuals will be selected from a random start on the list. The sampling design was decided based on the purpose of doing both the pilot and bridge study in one data collection effort. For the pilot study, in order to conduct the psychometric and factor analyses, the rule of thumb is to have at least 10 respondents for every item on the survey. Assuming the survey will have about 50 items, we will need at least 500 total respondents answering all survey items to conduct these analyses. Additionally, for multilevel psychometric analysis, we need at least 30 sites with sufficient representation in each. The total number of individuals sampled exceeds the minimum number needed for a single-level analysis and meets the sample size thresholds required for a multilevel analysis.

For the bridge study, we computed composite scores and associated standard deviations from the 2012 Hospital Survey on Patient Safety Culture Comparative Database and used them to estimate the statistical power we would have to detect a three percentage point difference in scores given different sample sizes of hospitals. A sample size of 40 hospitals was found to generate expected power of at least 0.74 to detect a three percentage point difference between the original and revised surveys. The power analysis for each composite score was for a paired t‑test, where an individual measurement pair was a sampled hospital’s two composite scores—one based on the original survey and the other based on the 2.0 survey. We estimated the standard deviation across hospitals of the paired differences by assuming bridge-study hospital-level composite scores would have the same standard deviation across hospitals as those in the 2012 database and that the correlation across hospitals between original-survey composite scores and corresponding 2.0-survey composite scores was 0.75 or greater[[1]](#footnote-1). The estimated power varied by composite scores and was above 0.80 for 10 of the 12 composites. Because the number of individuals surveyed within each hospital will be either a census or 500 individuals the power calculations focused on variability due to the selection of hospitals rather than the randomization of individuals to original vs. 2.0 surveys.

**Table 1. Distribution by bed size and teaching status of 40 hospitals selected for pilot test and bridge study**

|  |  |  |  |
| --- | --- | --- | --- |
| **Hospital Bed Size**  **(# of beds)** | **Type of Hospital** | | **Total Bridge Study Hospitals** |
| **Teaching** | **Non-teaching** |
| Small (49-99) | 5 | 8 | 13 |
| Medium (100-399) | 7 | 10 | 17 |
| Large (400+) | 4 | 6 | 10 |
| Total | 16 | 24 | **40** |

**Table 2. Distribution by bed size and teaching status of 18,375 individuals\* surveyed within 40 hospitals selected for pilot test and bridge study**

|  |  |  |  |
| --- | --- | --- | --- |
| **Hospital Bed Size**  **(# of beds)** | **Type of Hospital** | | **Total Individuals Surveyed** |
| **Teaching** | **Non-teaching** |
| Small (49-99) | 5 x 375 = 1,875 | 8 x 375 = 3,000 | 5,875 |
| Medium (100-399) | 7 x 500 = 3,500 | 10 x 500 = 5,000 | 8,500 |
| Large (400 +) | 4 x 500 = 2,000 | 6 x 500 = 3,000 | 5,000 |
| Total | 7,375 | 11,000 | **18,375** |

\* Assuming a 50% response rate, 9,188 individuals with completed hospital surveys will be available for analysis purposes

## 2. Information Collection Procedures

Cognitive interviews will include these steps:

* Mailing, faxing or emailing the surveys to the individuals recruited
* Receiving completed surveys via fax, email, or mail
* Telephone interview with respondents to discuss responses

The pilot test and bridge study survey data collection will include these steps:

* Programming the surveys for Web-based data collection
* Emailing sampled hospital providers and staff to notify them of the survey
* Reminder emails to nonrespondents

## 3. Methods to Maximize Response Rate

Cognitive interview participants will receive cash remuneration ($200 for physicians, $125 for registered nurses and department managers, and $75 for other staff). Although we are not offering remuneration for the pilot test or bridge study, we are providing a feedback report to hospitals as an incentive, and we are following up with reminder emails to nonrespondents.

## 4. Tests of Procedures

The procedures for this specific project have not been subjected to testing. However, the contractor, Westat, has conducted many similar projects and is using well-established research methods with this project.

## 5. Statistical Consultation and Independent Review

Input from statistical analysts was used to develop the study design and plans for data analysis, including:

Naomi Dyer Yount, PhD, Senior Study Director, Westat, 301-610-8842

Joann Sorra, PhD, Associate Director, Westat, 301-294-3933

Pat Dean Brick, MA, Senior Survey Methodologist, Westat, 301-517-4196

Richard Sigman, MA, Senior Statistician, Westat, 240-453-2783

1. For the Handoffs and Transitions composite score, the computed power for detecting a three percentage point difference was 0.74 for a bridge sample containing 40 hospitals. In the 2012 database, the average and standard deviation over hospitals of this composite score are 47 percent and 10.94 percent, respectively. The power analysis assumed that over hospitals in the bridge study the average of this composite score was 47 percent for the original survey and 50 percent for the 2.0 survey. [↑](#footnote-ref-1)