

THE SUPPORTING STATEMENT

EVALUATING RETENTION IN THE NATIONAL HEALTH SERVICE CORPS (NHSC)

CONTENTS

B. STATISTICAL METHODS.....1

1. Respondent Universe and Sampling Methods.....1

2. Information Collection Procedures.....4

3. Methods to Maximize Response Rates.....5

4. Tests of Procedures.....6

5. Statistical Consultants.....7

B. STATISTICAL METHODS

1. Respondent Universe and Sampling Methods

To assure that the final sample represents a highly diverse cross-section of the target population and that the sample provides adequate representation of all clinicians participating in NHSC, we propose to divide the targeted clinicians into a total of 72 sampling strata defined jointly by program type (Scholars vs. Loan Repayors), location of NHSC sites (i.e., urban, rural, or frontier), and clinical discipline and medical specialty. The NHSC is particularly interested in estimating differences in retention between disciplines and between clinicians in rural vs. urban vs. frontier locations. Clinicians in the 2010 file will not yield rates of actual retention since most are still serving and relatively few will have left their service practices. The 2010 sample will principally provide data on anticipated retention, a proxy for retention. The 2005/2006 sample will yield data on actual retention; therefore, its clinicians will be oversampled relative to the 2010 cohort. The smaller number of rural clinicians in both the 2005/2006 and 2010 groups will be oversampled relative to the more numerous urban clinicians. All clinicians serving in frontier counties are to be surveyed, as the NHSC is particularly interested in these individuals. A total of 7,464 clinicians and administrators from the 1998, 2005, 2006 and 2010 will be selected for the sample and our expected respondent sample size is 4,853. This assumes a 65% response rate, which is well above the norm for studies of clinicians but well within the response rates we have achieved in previous surveys of clinicians. The study team achieved a response rate of 75.2% for then-current NHSC clinicians when the previous, 1998 survey was conducted: the current survey is a follow-up to that survey. Using a combination now of both on-line and mail survey approaches, with better contact information through the improved NHSC data system, and a shorter survey instrument, we expect to match or exceed that rate now. The response rate in 1998 for NHSC clinician alumni was 58.9% and for site administrators was 62.1%. Again, with better survey methods, better contact information and a shorter questionnaire, we anticipate higher response rates now, with 75% the goal. Given their relatively small numbers, all respondents to the 1998 survey will be surveyed now.

We estimate the same, 65% response rate for administrators because (a) we achieved a 62.1% response rate from administrators in our 1998 survey and our survey methods will be better this time, and (b) we are only surveying administrators of practices where NHSC clinicians currently serve and many of these administrators are hoping to attract other NHSC clinicians in the near future, so they will have strong motivation to respond.

Restating our power calculations in more statistical terms, we will stratify our clinician sample by program type (Scholarship vs. Loan Repayment), location of NHSC service site (i.e., urban, rural, or frontier), and clinical discipline and medical specialty. Power calculations indicate that the number of respondents in groups to be compared should be at least 133 to detect a 15% difference in retention proportions (e.g., at 1 year, 4 years, etc.) between disciplines and specialties with the same size. Assuming a conservative 65% response rate and a 7% bad-address rate, cell size should be at least 220 to achieve a power of .80. Accordingly, any cells with fewer than 220 individuals will be surveyed in entirety. A total of 220 clinicians would be randomly selected from any cell where counts are greater than 220. The potential deficiency in this

approach is that we will have greater precision in estimates for rural clinicians than urban clinicians. We address this issue with rural vs. urban comparisons with larger, grouped samples, e.g., for all physicians versus all mental health disciplines, to provide ample group sizes.

	Frame Count N_h	Selected Sample Size n_h	Conservatively Estimated Respondent Sample Size $n_h * .65$
2010 clinicians			
Scholars	536	536	348
Frontier Loan Repayors	264	264	172
Urban Loan Repayors	3,670	1,556	1,011
Rural Loan Repayors	1,787	920	598
2005/2006 clinicians (recent alumni)			
Scholars	435	435	283
Frontier Loan Repayors	161	161	105
Urban Loan Repayors	1,908	1,366	888
Rural Loan Repayors	1,048	866	563
1998 clinicians (remote alumni)			
Respondents to 1998 survey	860	860	559
2010 Administrators			
Administrators of sites w/ NHSC clinicians	3,782	500	325
GRAND TOTAL	14,451	7,464	4,853

If the study encounters unforeseen difficulties achieving a 65% response rate, analyses would still be useful with respondent data obtained. A 50% response rate with a 7% bad address rate would still give us a power of .72 to detect a .15 difference in retention proportions for two groups, e.g., retention at 4 years of 25% of one group vs. 40% for another group. A 40% response rate and 7% bad address rate would provide a power of .62 to detect a similar .15 difference in retention proportions between two groups.

Power calculations for the administrator questionnaire indicate that a 65% response rate from the 500 administrators would give us a power of .94 to detect a 10% difference (45% agree vs. 55% disagree) between groups on our key outcome variables, which are Likert-scaled perception questions, e.g., administrators' perceptions that "Retention of NHSC physicians is more of a problem for our organization than retention of non-NHSC physicians" and "Retention of NHSC mental health practitioners is more of a problem at our organization than retention of non-NHSC mental health clinicians". Assuming just a 50% response rate, a power of .87 would be achieved to detect the same 10% difference between groups. Similarly, assuming a 40% response rate, a power of .78 would be achieved to detect a 10% difference.

Should this study's response rate fall well short of 65% (e.g., 50% or 40%), we will approach analyses in the following ways: (a) large between-group differences will remain statistically different and will still be identified in the data and through significant p-values, (b) when meaningful point-estimate differences between groups are noted but don't reach statistical significance, these will be reported as potentially real but with the caution that the differences did not reach statistical significance perhaps due to small group sizes, and (c) some of the planned analyses will be adjusted so that smaller groups are combined to reach sizes needed for adequate power in analyses, e.g., retention for clinical social workers and licensed professional counselors serving in rural areas may need to be reported for these two disciplines combined rather than for each group separately.

Again, the discussion above about how we will deal with a low response rate is provided for completeness and to demonstrate that this evaluation can go forward with meaningful and useful analyses even if subjects do not respond as anticipated. But we remain confident that the response rate will be strong—well over 50%, targeted to 65%, and likely to reach 75%. This is not a general population survey for which good response rates are now so difficult. This is a survey of highly educated, engaged and generally altruistic clinicians who are or have participated in a program that provided them from \$40,000 to \$200,000+ in support. And the survey is coming from a known and trusted source. Current and alumni NHSC clinicians are a highly motivated respondent group, as they have proven in our previous studies of the past 20+ years.

Completeness and Accuracy of the BMISS file. The BMISS file is the BCRS' active administrative file for contacting and monitoring its current clinicians, their contracts, their sites of service, etc. The BCRS also uses the file's information on its current clinicians to generate its workforce reports to Congress and others. The file is continually updated, with an expected lag of only 1 to 2 weeks for clinicians with new contracts to be added to the file. For these reasons, and having confirmed this with the staff who maintain this file, it is regarded as highly accurate, complete and current for its information on current clinicians. File information about practices where current NHSC clinicians serve is generally accurate for addresses, type of site, etc., but less so for the names of site administrators who we will be surveying. Therefore, administrator names and contact information were confirmed through calls to clinics (see 2. below).

The completeness of the BMISS files for identifying clinicians who served in 2005 and 2006 is also regarded by the file's creators and current staff as excellent. This information is simply the

retained administrative information that was maintained on these clinicians when they served. The contact information for these recent alumni, however, is not routinely updated and for many will be out of date. We therefore confirmed and updated contact information through an outside vendor (see 2. below).

The completeness of our list of respondents to the 1998 survey (remote alumni cohort) is 100%, as investigators of the current project carried out that earlier project and we now generated the list of these individuals from the earlier study's original database. No efforts had been made since 2000 to keep their contact information up to date, so current contact information was obtained from an outside vendor (see 2. below).

2. Information Collection Procedures

The sampling frame was identified through the NHSC's BMISS and legacy file systems of its clinicians and practice sites, as described above. NHSC's BMISS file also provides the starting point for respondent contact information. The BMISS file has been downloaded and analyzed within the work of this project (as discussed above). Contact information has been extracted from BMISS for subjects selected by the sampling approach. Contact information for the recent alumni (2005/06) sample has been verified and updated by sending the name/address list to the outside address verification firm. For participants in the previous NHSC survey 12 years ago where it is known that the contact information of the BMISS and legacy files is not current, updated contact information has been sought by matching against current health professions databases such as the CMS NPI and also through the same outside address verification firm. The initial source of contact information for current site administrators was also BMISS. This information was validated against HRSA's NHSC website of its qualifying practices and also by calling the selected sites to confirm/determine current contact information.

An e-mail will be sent to subjects containing a link to a web-based survey instrument which subjects are to click on to complete. Rejected emails and other indications of bad addresses will be sent to the external address finder service. After this approach has run its course and some subjects have not been reached, staff will attempt contacting the targeted respondents by calling the most recent NHSC site where the clinicians worked to request information on current addresses, etc. Sometimes we will learn of a state where a clinician moved to, which makes on-line licensure files the next source for contact information. On-line rosters of the various national professional associations, e.g., American Academy of Family Physicians, also provide information through which clinicians can be located.

As noted, some respondents may not wish to complete an on-line survey, and these individuals will immediately be mailed paper instruments to complete. Project staff will closely monitor responses. As strata fall behind expected response rates, staff will immediately follow-up with the non-respondents to promote participation through the approaches described in 3, below.

As responses are received the data will be directly loaded into a database for those responses received electronically. Responses received in hard-copy format will be keyed, edited and placed in the database. Keyed data will be keyed twice (by separate persons) and the resulting data

compared, inconsistencies resolved through agreement, and then placed into the master database. Item-specific responses in the database will be examined for completeness and consistency – the latter will be verified by considering cross item reliability, e.g., if a respondent indicates s/he is not married, then the items related to spouse (or partner) should be stated as NA, or be blank.

With regard to site administrators, the instruments tend to request subjective information based upon experience rather than precise, participant-specific information. The intent of the survey is to determine how best to retain valued personnel resources in underserved settings. Hence, the survey design (as it relates to administrators) focuses on best practices that have been found to be productive in retaining personnel as well as procedures, lack of benefits, etc., that have been found to inhibit retention.

Retention in our analyses will be defined several ways, e.g., proportions retained or anticipating retention at 12 and 48 months after service term completion, as well as proportions retained over time reflected as a continuous variable and analyzed with Cox Proportional Hazards Models. Retention will also be defined with respect to various settings, e.g., within clinicians' specific service sites, in practices within defined HPSA's for those who relocate to other practices, and in practice within any rural county. As stated above, survey data from the expected respondent numbers will allow us to report and compare retention (by the preceding various definitions): (a) among current and recent alumni participants of the Loan Repayment Program for those serving in rural versus urban versus frontier areas; for each discipline within urban areas and most disciplines within rural areas; for the largest disciplines within frontier counties; and for largest physician specialties within rural and urban areas; (b) current clinicians versus recent alumni for these same groups in (a); (c) among those in the remote alumni group (serving as of 1998) for physicians versus non-physician primary care providers versus dentists; rural versus urban participants; and (d) for those in the remote alumni vs. recent alumni versus current clinician groups the participants of the same disciplines as in (c); for those serving in rural versus urban versus frontier areas; and for the largest disciplines within rural and urban areas.

3. Methods to Maximize Response Rates

The proposed survey will use a variety of data collection strategies that seek to improve response rates as well as minimize respondent burden. Since the great majority of the clinicians are relatively young and technology-oriented, the survey team proposes to send the initial inquiries via email. The BCRS's data system, BMISS, contains the email addresses for the vast majority of current clinicians and for many clinicians serving in 2005 and 2006. For clinicians that now have different email addresses from those in the data system and for clinicians without an email address on the system, the team will obtain current email and physical address information from a service company that provides this information. The email invitation to subjects will offer those not wanting to respond via the Internet to either download the forms or to be sent paper copies of the questionnaire to then be returned by U.S. Mail. Also, those for whom no current email address on the BMISS file and for whom none is retrieved via the service company, we will call their practice sites to attempt to obtain a current email address. In cases where none of these approaches work, we will mail paper copies of the survey instrument to practice addresses.

Those who do not respond or those for whom the survey materials are returned as undeliverable, a search will be conducted for the person via the service company indicated. This company will not have current and accurate email and US Mail addresses for all subjects; therefore when needed, project staff will seek this location information through telephone calls to practice sites where these individuals are or served while in the NHSC.

As in all of our past studies, survey returns will be monitored continuously during the data collection phase, with returns continuously broken down by the study's key study groups, e.g., by cohort (current, recent, remote), discipline, specialty, and rural/urban/frontier location. Cross tab analyses will be run to compare different response rates across groups throughout the data collection phase. As certain groups prove slower to respond, cover letters in subsequent surveys will be tailored to better fit these groups. If reasons for slower response are not clear, we will call a small number of subjects from these groups to learn how they have perceived the study and about any hesitations to respond. We will then amend the cover letters and our other approaches to address these issues. If still meaningfully fewer subjects in certain groups are not responding, additional efforts will be made to reach them, i.e., an additional mailing and/or personal phone call from project staff requesting their participation.

Upon completion of the survey, non-response adjustment of the weights would be used to minimize response bias. This would be done by the calculating of response rate per stratum and multiplying each weight by the inverse of that response rate. This accounts for the uneven responses by frame.

4. Tests of Procedures

Virtually every item—including all key items—used in this project's questionnaires is drawn from questionnaires used by the project investigators in many previous studies. Repeated experience has demonstrated that clinicians understand these items, will respond to them appropriately, and find the items non-burdensome. The item wording and response options and formats have proven useful for analyses and presentation in reports and for dissemination. We have formally assessed and published the predictive validity of one of the instruments' central items, which asks clinicians how much longer they intend to remain in their current practices, their communities, and in settings that provide care to the needy. [Pathman DE, Konrad TR, Agnew CR. Predictive accuracy of rural physicians' stated retention plans. *Journal of Rural Health*. 2003;19:236-243.]

The survey approaches to be used to reach this study's clinicians and administrators have also been used extensively by the project staff. These methods have repeatedly been demonstrated to be successful in reaching clinicians, who can sometimes be hard to find, and obtaining high response rates. Key is to use multiple approaches to finding individuals (including the extra work of individual phone calls to current and previous employers), offering the option for subjects to respond electronically or by mail, and the use of repeated mailings.

On the questionnaire for administrators, they will be asked to report how long they have worked in their organizations. This is a standard question in surveys where individuals serving in

targeted jobs are asked to respond with information they will have learned through their work. If data from this question characterizes a respondent group as highly experienced in their organizations and/or roles, it helps establish their credibility and the accuracy of the information they report. Conversely, if responses indicate that many respondents are new in their organizations, then it suggests less credence should be given to their opinions, and in our study especially to respondents' opinions about NHSC clinicians who predated their work in their organizations. We also anticipate comparing perceptions of the NHSC and its clinicians from administrators with longer versus shorter tenures in their organizations.

The draft administrator questionnaire submitted with this project's OMB application had ambiguous wording to item 11b. "The availability of locum tenens is a key factor in the retention of NHSC clinicians in our organization.", to which possible responses range from "strongly agree" to "strongly disagree." For clarity, the wording for this item will be changed to "The ability of clinicians to get away for vacations and continuing education is key to the retention of NHSC clinicians in our organization."

5. Statistical Consultants

William D. Kalsbeek, PhD
Professor of Biostatistics
UNC School of Public Health
Director, UNC Survey Research Unit
919-962-3249

Guangya Liu
Research Assistant
Survey Research Unit
University of North Carolina, Chapel Hill
919-594-8925

Qi Xue, PhD
Senior Statistician
Quality Resource Systems, Inc.
703-352-7393