CENTER *for* STUDYING

HEALTH

SYSTEM

CHANGE

Community Tracking Study
Physician Survey
Round 1
Survey Methodology Report

Technical Publication No.

9

October 1998

Community Tracking Study
Physician Survey
Survey Methodology Report
(Round One, Release 1)

Center for Studying Health System Change

Prepared by:
The Gallup Organization
Government and Education Division
One Church Street, Suite 900
Rockville, MD 20850

Authors:
Linda Keil, Ph.D.; Manas Chattopadhyay, Ph.D.; The Gallup Organization
Frank Potter, Ph.D.; MPR, Princeton, NJ
Marie C. Reed, M.H.S.; Center for Studying Health System Change

Table of Contents

Tables

Community Tracking Study Physician Survey
Round 1 Methodology Report

1. Introduction

The Community Tracking Study Physician Survey is part of the Community Tracking Study, a comprehensive examination of the nation's health care system funded by the Robert Wood Johnson Foundation and conducted by the Center for Studying Health System Change (HSC)[1]. The purpose of the physician survey is to document changes physicians are experiencing in the health care system and to learn how these changes are affecting physicians, their practices and the way they deliver medical care to their patients. The goal is to provide information to public and private leaders that will enable them to make better policy decisions. The physician survey will be repeated every two years to track change over time.

A nationally representative sample of physicians was drawn from records maintained by the American Medical Association and American Osteopathic Association. Consistent with the overall design of the Community Tracking Study, physicians were sampled in 60 randomly selected communities across the United States. A separate random sample of physicians representative of the U.S. was also drawn to permit national tracking with greater precision.

The survey population includes physicians practicing in the continental United States who provide direct patient care for at least 20 hours a week and who are not Federal employees. Residents and fellows, as well as physicians in selected specialties are excluded. Primary care physicians were oversampled to permit analysis of certain aspects of their practice of medicine.

---

[1] For a discussion of the Community Tracking Study see Kemper, Peter et al, "The Design of the Community Tracking Study: A Longitudnial Study of Health System Change and Its effects on People," *Inquiry* 33: 195-206 (Summer 1996)

Under subcontract to HSC, The Gallup Organization conducted 12,385 computer-assisted telephone interviews (CATI) with a nationally representative sample of physicians. The interviews were conducted between August 14, 1996 and August 10, 1997. The overall response rate was 65.4%. To achieve this response rate, intensive tracing and refusal conversion efforts were required as well as persistent calling of respondents to gain cooperation. This report describes the sampling procedures, data collection methods, and weighting procedures in detail.

2. Sample Implementation

The Community Tracking Study Physician Survey sample is comprised of two independently drawn samples: a site sample and a supplemental sample. Each is a stratified sample designed to be representative of the nation. The supplemental sample permits greater precision of national estimates when combined with the site sample. The sample design requires the population of physicians to be classified according to geographic location, physician specialty category, and source of information for each of the two samples[2].

This chapter first describes the categories used and issues encountered in classifying the physician population and determining eligibility for the survey, followed by the sample allocation plan and the sample draw.

2.1 Classifying the Population of Physicians

The physician survey samples were drawn from physician records in the American Medical Association (AMA) Masterfile and the American Osteopathic Association (AOA) membership file.

2.1.1 Geographic Definitions

Separate geographic definitions were developed for the two samples.

1) **Site Sample**. The population of physicians was restricted to those practicing within the 60 communities randomly selected in the first stage sampling process as representative of the continental United States for the Community Tracking Study. The sites are defined as sets of counties comprising particular metropolitan statistical areas (51 sites) or rural areas within particular states (9).

---

2) **Supplemental Sample**. The population of physicians in the supplemental sample included physicians in the 48 continental states. The states were divided into 10 geographic strata. The strata were defined to match those used by the AMA in their Socioeconomic Characteristics of Medical Practice Study as follows:

1. Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
2. New York
3. Delaware, New Jersey, Pennsylvania, West Virginia
4. District of Columbia, Georgia, Maryland, North Carolina, South Carolina, Virginia
5. Alabama, Florida, Kentucky, Mississippi, Tennessee
6. Arkansas, Louisiana, Missouri, Oklahoma, Texas
7. Indiana, Michigan, Ohio
8. Illinois, Iowa, Minnesota, Wisconsin
9. Arizona, Colorado, Idaho, Kansas, Montana, Nebraska, Nevada, New Mexico, North Dakota, South Dakota, Oregon, Utah, Wyoming, Washington
10. California

2.1.2   Physician Specialty Category

Physicians were categorized according to whether their specialty is considered primary care or non-primary care.

1) **Primary Care Physicians**. For sampling purposes, primary care physicians were defined as physicians in the following primary specialties:

Family Practice
General Practice
General Internal Medicine
Internal Medicine/Pediatrics
General Pediatrics

2) **Non-Primary Care Physicians**. Non-primary care physicians include physicians whose primary specialty is any of the remaining specialties (except those specialties that were explicitly excluded from the study, as described below).

2.2 Exclusions from the Physician Population

All primary care specialties (defined as family practice, general practice, general internal medicine, internal medicine/pediatrics, and general pediatrics), as well as the medical, pediatric and surgical subspecialties, obstetrics/gynecology, and psychiatry are included in the sample for the Community Tracking Study Physician Survey. Radiology, anesthesiology, pathology and certain other selected specialties were excluded from the sample for the survey. This was done because major portions of the survey focusing on physician-patient interaction and management/referral of patients by primary care and specialty physicians do not apply well to the typical practice of physicians in these excluded specialties. Appendix A contains lists of excluded physician specialties with the AMA and AOA specialty designator and label.

Several other groups of physicians were excluded for a variety of reasons. The excluded groups are listed below followed by the reason for their exclusion:

Graduates of foreign medical schools who are only temporarily licensed to practice in the United States were excluded because they are not part of the permanent U.S. physician population.

1) Inactive physicians were excluded because they are not currently providing direct patient care.

2) Physicians not practicing in the United States were excluded because they are not currently providing direct patient care within the United States.

3) Federal employees were excluded because it is likely that their experience of the current health care environment is very different from those in the private sector.

4) Residents, interns and fellows were excluded because they are considered to be still in training.

5) Physicians who are not office-based or hospital-based (teachers, administrators, researchers, etc.) were excluded because it was considered unlikely that they provide at least 20 hours of direct patient care per week.

6) The national sample of physicians was restricted to the continental U.S. For this reason, physicians practicing in Alaska and Hawaii were excluded.

7) The AMA was also asked to exclude osteopathic physicians (DOs) since the sample of DOs was to be provided directly by the AOA.

8) The AMA excluded physicians who had been randomly sampled for the Socioeconomic Characteristics of Medical Practice Study conducted by the AMA in 1996.

2.3 Issues in Defining the Population of Physicians

Several issues arose in the process of defining the population of eligible physicians. These issues are described in this section.

1) Socioeconomic Characteristics of Medical Practice sample - Physicians sampled for the 1996 Socioeconomic Characteristics of Medical Practice survey conducted by the AMA were included in the population counts. However, at the request of the AMA, these physicians were excluded from the sample draw so as not to overburden them with survey requests.

2) No contact cases - Approximately 2% of the physicians listed in the AMA Masterfile have requested that their names and contact information not be given out for surveys or other purposes. It was decided to include these physicians in the population counts and in the sample draw for weighting purposes even though we would not be able to contact them for interviews.

3) Licensure variable - The initial sample specifications excluded unlicensed physicians. However, licensing information is not consistently captured in the AMA and AOA records. Excluding physicians whose records did not indicate they were licensed resulted in a reduction in the eligible population by approximately 50%. Thus, the license indicator variable was not used to restrict the sample. While it seems highly unlikely that a significant number of unlicensed physicians are practicing in the United States, we are unable to distinguish licensed from unlicensed physicians in our sample. Similarly, the AMA does not use the license variable in drawing its sample for the Socioeconomic Characteristics of Medical Practice survey.

4) Preferred Mailing Address - For purposes of assigning physicians to sites and to geographic strata, the Preferred Mailing Address on the AMA and AOA files was used. A potential problem with this variable is that the physician may prefer to receive mail at an address that is not his/her main practice location. For example, we know that about 40% of physicians listed in the AMA Masterfile prefer to receive AMA mailings at their residence. While most physicians live and work within the same metropolitan statistical area, there are some for whom the Preferred Mailing Address is not within the same site as their main practice location. Less commonly, there are some physicians in the supplemental sample whose Preferred Mailing Address is not in the same geographic stratum as their

main practice location (e.g., physicians who live outside of New York state, but practice within New York City).

The AMA Masterfile contains a field for Office Address. However, according to AMA staff, the information is often missing and is not updated as often as the Preferred Mailing Address. AMA considers Office Address a far less reliable variable than Preferred Mailing Address. For this reason, the Preferred Mailing Address was used even though a percentage of physicians would be misclassified for sampling purposes in regard to geographic location. The exact locations of the physicians' practices were verified during the interviews and, for analysis purposes, were classified based on this more accurate information.

2.4 Population Counts

The AMA and AOA were asked to provide counts of physicians in each of the sample design cells for both the site sample and the supplemental sample. Thus, for the site sample, population counts were provided by each organization for 120 cells (60 sites x 2 types of physicians). For the supplemental sample, population counts were provided by each organization for 20 cells (10 geographic strata x 2 types of physicians).

Initial requests for population counts for the site sample were forwarded to the AMA and AOA in April 1996. Once the initial counts were received, they served as the basis for the development of the optimal sample allocation described in the next section. However, the AMA and AOA files are not static. They are updated frequently with the result that the population counts are not stable. In practical terms, what this means is that although the initial population counts were used to develop the optimal sample allocation plan prior to drawing the site sample, the population counts at the time of sampling are actually a more accurate description of the population from which the samples were drawn. Thus, the counts at the time of sampling served as the basis for calculating the base weights.

2.5 Sample Draw

Sample specifications were provided to AMA and AOA for their use in drawing the samples. The specifications first described the limitations on the eligible physician population, then described in detail how the samples were to be drawn.

2.5.1   Site Sample

The AMA and AOA each received datafiles showing the number of physicians to be sampled in each of the 120 cells of the site sample from their respective databases.

These organizations were asked to draw the samples as follows:

Separate samples were to be drawn for each site and for PCP and NPCP groups within site for a total of 120 site samples.

Samples were to be drawn as systematic random samples as follows: to draw a sample of size n from a population of size N, the sampling interval (k=[N/n]) is first determined. An integer (say l) between 1 and k would first be selected at random and then every kth unit is selected in the sample. Hence the final sample includes the following units {l, l+k,...l+(n-1)k}.

Primary care physicians were oversampled at a rate of approximately 2.5 to permit analyses of this significant subgroup of physicians.

2.5.2   Supplemental Sample

The supplemental sample of physicians was subject to the same restrictions in terms of exclusions described above. However, once these exclusions had been completed, the supplemental sample was to be drawn from the entire eligible population of physicians practicing in the continental United States. It was to be drawn without respect to the sites defined for the site samples. Thus, it was anticipated that a small number of overlapping cases would be drawn; that is, some physicians would be drawn in both the site and supplemental samples.

The supplemental sample was to be drawn as a stratified random sample with the number drawn in each of the 20 strata (10 geographic strata X 2 types of physician strata) to be

8

proportionate to the size of the cell. Thus, if one of the 20 cells contained 5% of the total

population, then 5% of the sample cases would be drawn from that cell.

There was no oversampling of primary care physicians in the supplemental sample. Primary

care and non-primary care physicians were drawn in proportion to their occurrence in the

population.

2.6 Sample Variables Requested from AMA and AOA

AMA and AOA were asked to provide specific variables from their files for each sampled

case. This section lists the variables provided by these organizations with brief descriptions

where necessary. Variables provided by AMA and AOA include:

- First name
- Middle name - Middle names and initials for AOA cases are included in the First name field
- Last name
- Suffix - Example: John Smith, III would have code 3; provided by AMA only
- Address - Field contains street address
- City
- State
- Zip code
- Telephone number - Approximately 10% of cases received did not have telephone number
- Medical school
- Year of graduation
- FIPS state and county code
- Primary Metropolitan Statistical Area/Metropolitan Statistical Area - not provided by AOA
- U.S. Department of Commerce Region - not provided by AOA
- U.S. Department of Commerce Division - not provided by AOA
- Year of birth
- Major professional activity
- Primary specialty
- Secondary specialty - not provided by AOA
- Present employment
- American Specialty boards (up to 3) - not provided by AOA
- ECFMG number (indicator of foreign medical school from which physician graduated) - not provided by AOA
- Type of practice
- Gender

- Training segments including training specialty, beginning and ending dates of training and training institution. AMA provided up to nine training segments per record. AOA did not provide this information.

2.7 Quality Control Steps in Sample Preparation

Once Gallup received the samples from AMA and AOA, the various files were checked, concatenated, and prepared for fielding. This section describes the steps that were taken to prepare the files for the field.

First, a variety of quality control checks were made on each of the files. The most important check was to verify that the number of cases requested in each of the sample design cells was consistent with the number of cases received.

Next, the various files were concatenated by matching similar fields. So, for example, the AMA variable SCHOOL and the AOA variable COLLEGE were considered the same for the merge. In the process of merging, some information provided by the AOA was left out. AOA had provided a FAX telephone number. Since there were so few of these available and since AMA had not provided the information, it was not included in the concatenated file.

One variable provided by AOA, but not by AMA, was included, however. The variable is AOATYPE. It is missing for all AMA cases. For all AOA cases, however, if AOATYPE=0 the case is from the supplemental sample, if AOATYPE=1 the case is a primary care physician, and if AOATYPE=2 the case is a non-primary care physician.

After merging the sample files into one large file, a de-duplication program was run. This program identified the cases that had been sampled in both the supplemental and site samples. These cases were removed from the main sample file and kept in a separate file. The de-duplication process was not as straightforward as it may sound. Many people have the same names especially if the names are common. This is compounded among physicians by a

tendency toward medical families especially with sons following their fathers into medical practice. In some cases, it is not possible to distinguish between individuals even when matching their full address and telephone numbers in addition to their full names. The AMA includes a variable called SUFFIX as part of the set of name variables. This variable distinguishes between Sr. and Jr., for example. Gallop's de-duplication process included matching first, middle, and last names, suffix, birth year, and zip code. Although AMA had been asked to exclude DOs from its sample draw, we checked for duplicate cases that might have been erroneously selected by both AMA and AOA. No such duplicates were found.

Prior to fielding the sample, Gallup added several control variables to the file. The main control variable is the Gallup Identification number. This is a five-digit number that is unique to each case. All potentially active cases received a Gallup ID number. Cases identified by the AMA as "No Contact" cases did not receive a Gallup ID number because we were not permitted to field these cases. All cases including the "No Contact" cases received a site code from 00 (Supplemental) to 60.

The AMA and AOA primary specialty codes were recoded from alphabetic codes to three digit numeric codes. These three digit specialty codes controlled the case through the interview with regard to whether or not it received the primary care questions. All cases started with a specialty code designated by either AMA or AOA. (However, during the course of the interview, a series of questions permitted the respondent to confirm or change this primary specialty designation. Of those who completed interviews, 7.4% changed their primary specialty designation while 92.6% agreed with the AMA/AOA designation.)

The final step before the field period began was to replicate the sample. Replicates were formed by randomly selecting 67 cases for each replicate without replacement. The entire sample file including the "No Contact" cases was first sorted by site number. The number of cases in

each replicate was 67. To form replicates, each case in the sorted file was numbered from 1 to 67. A total of 432 replicates were formed in this fashion. The method ensures proportional representation of all sites in the replicate structure.

Every case in the sample file was assigned to a replicate including the "No Contact" cases. Since we did not anticipate releasing the entire available sample, it was important to assign replicate numbers to the "No Contact" cases so it could be determined whether or not the case would have been released when calculating weights.

As it became apparent that certain cells (defined by site and PCP/NPCP) would reach their target numbers of completes without further release of sample, releases were stopped in these cells. Table 1 shows how the sample was released. There are variations across replicate release groups in the number released because of the differential number of "No Contact" cases and, in later replicate release groups, because we stopped releasing cases in certain cells.

<div align="center">Table 1</div>
<div align="center">**Sample Released**</div>

Replicate Release Group	Replicates Included	Number of Cases Fielded	Number of No Contact Cases	Number of Cases Released	Date Released
#1	1-42	2,733	79	2,812	8/15/96
#2	43-84	2,763	50	2,813	9/2/96
#3	85-105	1,371	36	1,407	9/9/96
#4	106-126	1,369	36	1,405	9/16/96
#5	127-147	1,377	28	1,405	9/23/96
#6	148-189	2,763	48	2,811	10/7/96
#7	190-210	1,382	25	1,407	10/15/96
#8	211-252	2,747	65	2,812	10/21/96
#9	253-294	2,541	62	2,603	11/1/96
#10	295-315	1,231	32	1,263	11/9/96
#11	316-345	1,590	32	1,622	11/15/96
#12	346-432	1,229	32	1,261	2/24/97
TOTALS		**23,096**	**525**	**23,621**	

In Replicate Release Groups 1-8, cases were released in all sites. However, restrictions were placed on the sample release beginning with Replicate Release Group #9, for reasons discussed below.

Physician Location Errors, Specialty Classification Errors, and the Final Sample Release. A preliminary dataset was prepared for analysis that contained roughly the first 4,000 interviews completed. One aspect of the analysis of this preliminary data involved comparing the location and specialty of the physician when sampled with the physician's location and specialty at the time of interview. Physicians whose main practice location at the time of interviewer was in a different site than where they were sampled are called "Movers." In reality, probably few physically moved. In many cases, it appeared that the sampling address (from the AMA or AOA frame) was the physician's home address, but he/she actually practices in another site. Nevertheless, to simplify discussion of such cases, they are called "Movers."

Since we had no systematic way to capture physicians moving into a site, the only thing we could look at was the number of physicians moving out of the site. It was found that in some sites a relatively large proportion of physicians were misclassified on the basis of their selection address, that is, they were "Movers" out of the sites. Among the high intensity sites, Orange County was of particular concern because 14.6% of physicians sampled in the site were found to be "Movers" at the time of interview. Similarly, Newark lost just over 10% because of location errors.

In like fashion, the physician's primary specialty at the time of interview was compared with specialty used for sample selection. As noted earlier, 7.4% of physicians interviewed gave a primary specialty different from that listed for them by AMA or AOA. Specialty classification errors of particular concern were those that changed the physician's designation as PCP or NPCP. Among physicians sampled in the PCP stratum, 5.8% (n=731) were found to be NPCP when interviewed. Conversely, 1.6% (n=200) of those sampled as NPCP were found to be PCP.

Evaluation of the impact of these location errors and specialty classification errors on the final numbers of completed cases per site led us to become concerned about having a sufficient number of cases in certain cells of the sample design, particularly in the high intensity sites and among primary care physicians. The result of this concern was that in February 1997 one last set of cases was released to the field. This final release group included only primary care physicians in high intensity sites. The purpose of the extra release of these cases was to improve the final number of completed interviews in these cells of the sample design where the impact of location error and specialty classification error was relatively large.

3. Survey Design and Preparation

3.1 Schedule

Gallup was awarded the contract to conduct the Community Tracking Study Physician Survey on March 1, 1996. The following table shows the dates for key activities.

Dates	Activities
3/1/96	Contract award
3/4/96 - 7/25/96	Sample design, sample specifcation development, sample Draws by AMA and AOA
3/4/96 - 5/3/96	Questionaire revisions/Interviewer training materials developme nt
5/1/96 - 7/10/96	Obtain study endorsements/Prepare advance leter
5/6/96 - 5/7/96	Pilot test interview training
5/8/96 - 5/20/96	First pilot test
6/7/96	Second round pilot test interviewer training
6/11/96 - 6/17/96	Second pilot test
7/22/96 - 7/23/96	Main study interviewer training
7/24/96 - 7/31/96	Interviewers conduct practice interviews
7/25/96	Prepare sample for field
8/9/96	Mailout of advance letters
8/14/96 - 8/10/96	Conduct interviews
11/15/96	1st Data Delivery
3/10/97	2nd Data Delivery
8/19/97	Final Data Delivery

3.2 Instrument Development

HSC contracted with Project Hope and CODA to develop the instrument and conduct cognitive testing. Beginning immediately after contract award, Gallup also participated in questionnaire discussions and revisions.

The questionnaire[3] contains 10 sections as follows:

Section A. **Physician Supply and Specialty Distribution (Questions S1-S13 and A1-A19)** - the S series questions capture tracing information in case the sampled physician is not at the telephone number given; questions in the A series establish eligibility, number of practices, location of primary practice, year began medical practice, primary specialty, board eligibility and certification, and current level of satisfaction with overall career in medicine.

Section B. **Physician Time Allocation** - number of weeks practiced medicine in 1995, hours worked during last complete week of work, hours spent in direct patient care during last complete week of work, hours in the last month spent in charity care, and (for physicians with more than one practice) percentage of direct patient care time spent in main practice.

Section C. **Practice Arrangements and Ownership** - respondent ownership of practice, type of practice, other owners of practice, number of physicians employed by practice, number of non-physician medical practitioners employed by practice, whether physician was part of a practice that was purchased by another practice or organization during the past two years.

Section D. **Gatekeeping / Medical Care Management Strategies / Scope of Care** - all physicians: level of effect that various medical care management techniques have on the physician's practice of medicine. PCPs only: percentage of patients for whom physician acts as gatekeeper, change in severity or complexity of patients' conditions for which care is provided without referral to specialists, appropriateness of complexity or severity of patients' conditions for which care is provided without referral to specialists. Specialists only: changes in complexity or severity of patients' conditions at time of referral by primary care physicians, appropriateness of complexity or severity of patients conditions at time of referral, change in number of referrals received.

Section E. **Practice Styles of Primary Care Physicians** - clinical descriptions of patient histories for which physician is asked to state the percentage of patients that he/she would refer, hospitalize, provide the treatment, etc.

Section F. **Ability to Provide Care / Ability to Obtain Needed Services for Patients / Acceptance of New Patients with Various Types of Insurance** - level of agreement with statements regarding having adequate time with patients, freedom to make clinical decisions, ability to provide high quality care, level of communications with specialists/primary care physicians, ability to maintain continuing relationships with patients, ability to obtain a variety of specified services for patients, acceptance of new patients insured by Medicare, Medicaid, private insurance.

---

[3] For specific survey questions, see *Community Tracking Study Physician Survey Instrument,* Washington, DC: Center for Studying Helath System Change, Technical Publication Number 3, September 1997.

Section G.   **Practice Revenue** - percentage of practice revenue from Medicare, Medicaid; number of managed care contracts; percentage of practice revenue: from managed care, paid on a capitated or other prepaid basis, from largest managed care contract; proportion of revenue from largest contract which is capitated or prepaid.

Section H. **Physician Compensation** - whether physician is salaried, eligible to earn bonus or incentive income, factors used by practice to determine compensation, percentage of 1995 income earned in the form of bonuses, returned withholds, or other incentive payments, amount of income in 1995.

Once agreement was reached on the content of the questionnaire, Gallup staff prepared it for pretesting. First, the questionnaire was typed in the Gallup CATI format which includes interviewer instructions as well as skip pattern instructions and column locations that are used by the CATI programmer. Attention was also given at this point to the response category codes. In particular, consistent codes were selected for missing values as follows:

8 (98; 998) = Don't know

9 (99; 999) = Refused

One particularly critical aspect of this technical editing process relates to physician specialties. As described in Chapter 2, the physician sample contains two strata: (1) primary care physicians (PCP) and (2) non-primary care physicians (NPCP). Additionally, a number of physician specialties were excluded from the study. (See Appendix A for complete specifications of excluded specialties.) In Section A of the questionnaire, physicians were asked to confirm their primary specialty. Some physicians' primary specialties were different from those listed in their AMA or AOA records. Occasionally, this meant the physician was not eligible for interview because his/her primary specialty was among those excluded from the survey (a total of 66 physicians were excluded from the survey for this reason). Sometimes it meant the physician had to be reclassified from PCP to NPCP or the reverse. And, sometimes, it made no difference with regard to the sampling stratum. Part of Gallup's task during technical editing was to work with HSC to develop the proper set of exclusion and reclassification rules based on physician specialty.

Many of the patterns throughout the survey were determined by the physician's stratum. In particular, the primary care physicians were asked Section E, the practice style questions, but non-primary care physicians were not asked these questions.

3.3 Pretest

The purpose of the pretest of the Community Tracking Study Physician Survey was to test the questionnaire in terms of skip patterns, wording, and other content factors. Another important goal was to evaluate the time required to administer the interview.

Pretest Sample. The sample of physicians for the pretest was drawn from the AMA Masterfile. Both MDs and DOs were included in the sample. Physicians were sampled from outside of the 60 study sample sites with oversampling in California and Minnesota. These states were oversampled because of their higher penetration of managed care. We were hoping to be able to interview several physicians practicing in HMO settings as a test of the questionnaire's skip patterns and wording. Eighty percent (80%) of the sampled physicians were primary care physicians (PCP) and 20% were non-primary care physicians (NPCP). The goal of the pretest was to complete 50 interviews, 40 with PCPs and 10 with NPCPs.

Pretest Dates. Two sets of pretest interviews were completed. Considerable reprogramming of the instrument occurred between sets. The dates of these pretests were as follows:

Pretest #1: May 8 - May 20, 1996 (18 days) - 27 completed interviews; 16 PCP and 11 NPCP.

Pretest #2: June 11 - June 17, 1996 (7 days) - 25 completed interviews; 21 PCP and 4 NPCP.

Procedures. The pretests were intended to test only the interview itself, not the study procedures. Thus, no response rate estimates were computed for the pretests. The advance

materials that were used for the main study were not ready for the pretest. Instead of sending an advance letter from the Robert Wood Johnson Foundation as we did for the main study, a letter from Dr. Robert St. Peter of HSC was faxed to the respondent's office on request. This letter was very helpful in gaining cooperation and physicians generally responded positively to the study.

Length of Interview. It was clear from the first set of pretest interviews that the interview was too long. The average interview length in Pretest #1 was 31.1 minutes (34.0 for PCPs and 28.0 for NPCPs). Prior to Pretest #2, cuts were made to reduce the time required for the interview. Overall, these cuts succeeded in reducing the interview length by 6.5 minutes to 24.6 minutes (25.2 for PCP and 23.0 for NPCP).

The length of the interview was still of some concern following Pretest #2. A goal articulated by the questionnaire design team was to have 90% of the PCP interviews at or below 25 minutes. In Pretest #2, just 57% of PCP interviews met this criterion, with 15% of interviews greater than 30 minutes in length. Following the second pretest, additional questions were cut from the interview to meet the goal of 90% of PCP interviews at 25 minutes or less.

Some reasons for shortening the interview are the following. While the interview generally seemed to be interesting to physician respondents, it was also cognitively burdensome in some places. For example, many of the questions were long because they included explanations and definitions to convey the meaning of the question. In listening to the pretest interviews, we felt strongly that the cognitive burden was eased when interviewers slowed down. Yet, doctors are busy and interviewers are often pressured to speed through the interview. The longer the interview, the more likely it is that interviewers would be pressured to speed up with a resulting increase in cognitive burden and potential decline in data quality.

Formatting and Other Changes. In addition to cutting questions, several of the question wordings were revised after reviewing pretest respondents' reactions. A number of format changes and other minor changes were also made following the pretests. For example, changes and corrections were made to the CATI programming, to response category wording, and to interviewer instructions appearing on screen.

3.4 Advance Letter Preparation

Obtaining Physician Association Support. In order to maximize physician interest and participation in the survey, major physician associations were solicited for their support of the survey. Gallup contacted the agencies, described the project, and obtained the support of all but one organization. Associations differed considerably with regard to their processing of this request. In some associations, a formal letter of request that summarized the project was sufficient. In others, the request was reviewed by their boards and required quite a lengthy process. The names of the supporting organizations are included in the advance letter sent to physicians, which can be found in Appendix B.

Mailing of Advance Letters. An advance letter was prepared and mailed to sampled respondents one week prior to the release of the sample to interviewers for calling. In addition to the letter describing the survey and asking for the physician's participation, the mailing included a copy of a brochure describing the Center for Studying Health System Change (www.hschange.com.about.html). Since the sample was released in waves (see Chapter 2), waves of advance letters were sent one week prior to each release.

A second copy of the advance letter was also sent to many respondents at different times throughout the field period. For example, refusal cases were permitted to "age" for a period of time, then were assigned to refusal specialists for attempted conversion. One week prior to assignment of refusal cases, a second copy of the advance letter was sent. In November, 1996,

the AMA Newsletter published an article describing the importance of the Community Tracking Study Physician Survey. After publication, a copy of the article was also included when advance letters were mailed to respondents.

During the project's seventh month in the field, a second letter was prepared and sent. A copy is shown in Appendix B. This letter emphasized that the data collection would be ending soon and provided an advance incentive check to encourage speedy response. This strategy is discussed in greater detail in Section 4.8.

3.5 Computer-Assisted Telephone Interviewing System (CATI)

Gallup uses the SURVENT CATI system. Gallup's SURVENT programmers prepared the instrument for CATI administration of both pretests as well as for the main study. Following each of the pretests, changes were identified and made in the program. Gallup's proofing department carefully proofed each of the CATI screens and extensively tested the instrument to be sure the program was working as intended. Instrument testing was also conducted by project staff familiar with content issues.

3.6 Telephone Management System (TMS)

SURVENT interfaces with the Telephone Management System (TMS). The TMS is an automated sample server that distributes telephone numbers to each interviewer according to the sample design. It maintains call histories on every released case to support reporting including call statistics and interviewer productivity figures.

The physician survey was an executive ownership study. This means the study was conducted by Gallup's executive interviewers who specialize in interviewing physicians, other health professionals, and business executives. Executive ownership also means the interviewers

"owned" their cases. Interviewers were responsible for setting and keeping their own callback appointments just as in-person interviewers do in the field. They had ample opportunity to establish rapport with office workers and the physicians themselves when calling to set appointments and conduct interviews. The TMS system keeps track of which cases are owned by each interviewer and collects extensive histories of each call attempt for every case.

3.7 Interviewer Selection

Gallup's executive interviewers conducted the physician interviews. These individuals are career interviewers whose experience ranges from 3 to 15 plus years. All are full-time Gallup employees. This executive interviewing team devotes a substantial proportion of its time to studies of physicians and other health professionals, completing approximately 25,000 physician interviews per year.

Gallup's original plan had been to train 30 interviewers to work on the physician survey. However, because of delays in sampling and questionnaire development, it was decided that a larger team was needed to complete the study within the planned field period. Therefore, a total of 49 executive interviewers were trained to conduct the physician survey interviews.

Partway through the field period, it was discovered that a number of cases, especially in the Miami study site, could not be completed because either the physician or his/her receptionist spoke only Spanish. To solve this problem, Gallup trained one additional interviewer who is bilingual in English and Spanish to take over these cases. Thus, a total of 50 Gallup interviewers worked on the physician survey.

3.8 Interviewer Training

Three interviewers were trained to conduct the pretest interviews. This preliminary training provided the opportunity to test the interviewer training materials and agenda. Revised materials and training agenda were then prepared for the main study training.

Training Materials. Interviewers were encouraged to keep these materials in their carrels when making Physician survey calls. Following is a list of the interviewer's materials:

1) Physician specialty list -alphabetical listing
2) Physician specialty list- list by categories
3) Definitions of key terms
4) Copy of Robert Wood Johnson Foundation advance letter
5) Copy of brochure describing the Center for Studying Health System Change
6) Interviewer's manual[4].

Training Agenda. The training consisted of lectures providing background about the study, the sample of physicians, and the interviewers' role on the project. Discussions of gaining respondent cooperation were led by the pretest interviewers on the basis of their experience with the study introduction and supporting materials. Introduction to the survey itself was conducted on-line. Interviewers took turns reading questions and gained experience with all aspects of the instrument through the course of several passes through the survey following different skip patterns.

Exercises were prepared for interviewers to complete at home after the first day of training. The correct answers were provided and discussed on the second day. The exercises were designed to emphasize key points regarding respondent eligibility to help interviewers become familiar with issues leading to the important eligibility decision.

---

[4] *Community Tracking Study Physician Survey Interviewer Training Manual*, Washington, DC: Center for Studying Health System Change, Technical Publication Number 6, July 1998.

Interviewers were also able to practice the interview by working in pairs to conduct four mock interviews that had been prepared to illustrate particularly difficult aspects of the interview. Following the completion of the training sessions, interviewers were required to conduct a mock interview of the project director or assistant project director. These interviews were conducted by telephone and were evaluated so interviewers received feedback on their mastery of the interview prior to beginning actual interviews.

3.9 Preparing Sample for the Field

Once the sample file was prepared and checked as described in Chapter 2, the combined sample file was sent to a vendor for telephone number look-up. Although approximately 90% of the records contained a telephone number, it was known that the information in the AMA and AOA files might be out-of-date. So, we thought it prudent to attempt to obtain the most recent telephone number for each case possible.

The first step was to run an electronic look-up procedure assigning current telephone numbers to as many cases as possible. Cases for which a match could not be found in this way were sent on for manual look-up. This second step required calling directory assistance for the correct telephone number. A "tight match" process was used for both steps meaning that a telephone number was considered a match to a sampled physician only if the listing was for a person with exactly the same name, at the same address, in the same city and state as the sampled physician. If these conditions could not be met, then no new telephone number was assigned. For approximately 10% of the sample, no telephone number could be found either in the AMA or AOA records or through the telephone look-up process.

It should be noted that if AMA or AOA provided a telephone number, this number was available to the interviewers to try even if the look-up procedure had identified a number

considered more recent. Thus, in the majority of cases, at least two possible telephone numbers were available at the start of the study.

After the quality control checks had been run on the sample records, including the telephone number look-up procedures, the sample was ready to enter into the Telephone Management System. Records with no telephone numbers were separated from cases ready for the field. Cases without telephone numbers were assigned to the tracing team to locate the respondent and provide a current telephone number. Such cases were considered "released" into the study sample as their replicate groups were released even though their immediate status was "tracing."

4. Data Collection

Gallup executive interviewers spent 25,495 hours working on the Physician survey. During that time, they conducted 12,385 completed interviews. As noted earlier, 50 interviewers were trained to work on the study. The interviewers were divided into three groups (high, medium, and low) according to the number of hours they spent each week on the project. Gallup's most experienced physician interviewers were in the "high" and "medium" groups. They were responsible for 65% and 27% of the interviews, respectively. The "low" group, though experienced executive interviewers, have less experience conducting physician interviews than interviewers in the other two groups and conducted only 7% of the interviews.

4.1 Telephone Center Staff

In addition to the 50 executive interviewers, Gallup's telephone center staff assigned to the Physician survey included four supervisors (including the head supervisor of the Telephone Center), and support staff. Supervisory duties included monitoring interviews, reviewing and resolving problem cases, producing reports, and coordination of listen-ins by HSC. Support staff routed Physician survey calls coming in on the 800 line to appropriate interviewers, taped some interviews for in-depth evaluation by supervisors, and conducted validation calls.

4.2 Interviewer Monitor

Gallup's quality assurance program includes interviewer monitoring and interview validation. For this project, just over 11.5% of the interviewers' work was monitored. Two types of monitoring are included in this total. First, Gallup supervisors randomly tape interviews conducted by each interviewer for in-depth evaluation and performance feedback. About 2% of the interviewer's work on this project was monitored in this way. The remaining 9.5% was monitored in real time by supervisors who listened in with silent monitoring equipment to

26

interview attempts, refusal conversion attempts, and full interviews, and then provided feedback to interviewers to improve performance.

As an additional quality check, Gallup also validates at least 10% of completed interviews by calling respondents, asking if they completed the survey, and checking key items to make sure the entire survey was completed. For this study, 2,179 surveys were validated (17.5%).

4.3 Sample Release Strategy

As shown above, the sample was released in replicate groups ranging from 1,264 cases to 2,814 cases. The majority of the sample was released within the first three months of the field period with one final sample release in the sixth month. As cases were released, interviewers received case assignments appropriate to the number of hours they had committed to the project.

Our original idea had been to close out all cases within a replicate release group after three months in the field. This would have staggered the closing of cases and would, it was hoped, provide information about response rates and refusal conversion percentages that could have been used in decisions about how many additional cases to release to the field. This strategy did not work out as anticipated for several reasons. First, the number of cases requiring tracing was much higher than anticipated. The tracing efforts took longer than anticipated as well. Thus, closing replicates after a three-month field period would have lowered the number of tracing cases that were located with resulting adverse effects on the response rate. Another reason for keeping the replicates open throughout the entire field period instead of closing them on a staggered basis as planned was that refusal conversion efforts were more likely to be successful the longer the case had "aged" since the refusal. Attempting to convert refusals within one or two months of receiving the refusal was much more likely to lead to a final refusal than refusal conversion attempts four, five, or six months after the refusal.

4.4 Length of Interview

The average length of interviews for the Community Tracking Study Physician Survey was 19.4 minutes. In general, the interview was longer for primary care physicians than for non-primary care physicians because the former were asked to respond to the series of questions on practice style. The average length of interview for PCPs was 21.6 minutes while the average length of interview for NPCPs was 17.7 minutes. The number of incomplete interviews was very low, just 72 or 0.5% of the interviews started. Although interviewers sometimes had to schedule callbacks to complete partial interviews, it was rare that they were unable to do so.

4.5 Spanish-Speaking Physicians

In sites with sizable Hispanic populations, there were physicians who did not speak English well enough to complete the interview in English. In some cases, the initial interviewer could not get past the office worker to determine whether the physician him/herself spoke English. Often, in these cases, it was just a matter of having a bilingual interviewer call and talk with the gatekeeper, describe the study, and ascertain whether the physician could speak English. Then, the regular interviewer could complete the interview with the English-speaking physician even though his/her office staff spoke only Spanish.

However, 61 cases were eventually assigned to a bilingual interviewer. In all of these cases, the information we were able to obtain suggested the physician spoke only Spanish or did not speak English well enough to complete the interview. Unfortunately, these physicians were also extraordinarily difficult to interview. Of the 61 assigned to the bilingual interviewer, only 3 interviews were actually completed in Spanish. Most of the remaining 58 cases either refused or could not be contacted.

4.6 Tracing

The tracing effort required for this study was considerably greater than had been anticipated. Overall, 20.7% of all cases released to the field had to be traced, either because we did not have a telephone number or because the telephone number provided to the interviewer was incorrect. In this section, the tracing efforts are described in some detail.

As described earlier in Section 3.9, one step in preparing the sample for the field was to do a search for the most current telephone number. After the "tight match" procedure was completed, 2,356 cases (10.2% of the sample) had no telephone number either from the telephone number match procedure or from the original AMA/AOA sample records. These cases could not be fielded without a telephone number, so they immediately became tracing cases.

Of the cases that were sent to the field, 2,472 cases (10.5% of the sample) were returned for tracing because the telephone numbers were incorrect. Prior to returning a case for tracing, however, the interviewers attempted to obtain a correct number for the physician from the person who answered the phone at the number given. Interviewers also attempted to reach these respondents at the alternative telephone numbers provided by AMA/AOA when these were available. So, only those cases for which a good telephone number could not be obtained were returned for tracing.

Overall, Gallup's tracing staff located 43% of the tracing cases (n=2,077). We were more successful in tracing the cases that had no telephone number than cases returned for tracing: the find rates were 49% and 37% respectively. This is not surprising: interviewers often obtained forwarding information for cases that were fielded with incorrect numbers, so it was only the more difficult cases that were returned for tracing.

Normal Tracing Procedures. Gallup followed four basic tracing steps to locate respondents. The "tight match" procedure that was used to identify the most current telephone number is the first step and has already been described.

Loose Match Procedure. The second step, once a case was identified as either having no telephone number or was returned from the field because the telephone number provided was incorrect, was to send the case to the vendor again. This time, an attempt to identify a possible telephone number was made using a "loose match" procedure. In this procedure, variations in the first and middle names are permitted and it is not necessary to match the street address, only the city and state. The following are examples of tight and loose matches.

Case Information	Tight Match	Loose Match
Samuel W. Abrams	Samuel W. Abrams	S.W. Abrams
1234 W. Main Street	1234 W. Main Street	789 Broadway
Hometown, AL	Hometown, AL	Hometown, AL

Confirmatory Calls. When using the loose match procedure, it is necessary to follow up on any potential matches to verify that the telephone number provided is the correct number for the sampled respondent. Thus, the third step in Gallup's tracing process was to make confirmatory telephone calls to all of the cases for which the loose match procedure provided potential telephone numbers. Gallup interviewers confirmed the following information during these calls:

- Full name
- Primary specialty
- Birth date

If all three of these elements were correct, the case was considered found. Interviewers also colleacted the correct current address of the respondent during this call so an advance letter could be sent prior to attempting to complete the interview.

CD Searches. If the loose match procedure failed to provide a new telephone number, or if the number provided by this process proved to be incorrect, a series of CD searches were

performed. Gallup obtained a set of CDs containing all the white and yellow page listings in the U.S. There are seven CDs in the set. One CD contains all of the business listings for the entire U.S. The other six disks contain both residential and business listings in each of six regions.

The process of searching for possible matches on the CDs varied by case depending on the outcome of prior tracing steps and the information provided. In general, the procedure started with a search by the physician's name on the business disk. Since this covers the entire country, it was most efficient to begin with this disk. It was also possible to search the business disk by address alone. In some cases, it was possible to find a physician who worked for a large hospital or HMO and who was not listed individually by searching under the address.

If these searches failed, the next disk to be searched was the regional disk appropriate to the address in the case record. So, for example, when searching for a case sampled in Miami, we would first search the south regional disk. Again, the disk was searched first by the physician's name and then by the address alone.

In some cases, additional regional disks were searched. Whether this occurred or not depended in part on the outcome of the prior searches. If potential telephone numbers were found on the business disk, confirmatory calls would be made to these numbers before searching the regional disk. Similarly, if possible numbers were found on the first regional disk, they would be called before additional regional disks were searched.

Another limiting factor on searching was the nature of the physician's name. If the name was very unusual, it might be possible to find the physician by searching all six regional disks and calling any possible numbers uncovered. However, with very common names, this strategy was not feasible. Although we did not set an absolute limit on how many possible numbers could be called, we rarely called more than 10 numbers for any doctor. If the doctor had a common name,

we used other information to narrow the list of possible numbers to the most likely set. Occasionally, we had information that pointed our search in a particular direction. For example, on the initial call, the interviewer may have learned that the physician moved to a particular state even though the person who answered the telephone did not know the doctor's new address or telephone number. Armed with the state name, however, we could often locate the physician by searching the appropriate regional disk.

Special Tracing Procedures. In addition to the normal tracing procedures just described, special tracing procedures were employed for certain cases. This section describes these efforts.

Database Searches. Early in the field period, we investigated two alternatives for electronic telephone number searches. One alternative was to send tracing cases to the telephone number look-up vendor for their loose matching procedure. As described above, this is the alternative we ultimately used. However, as a test, we initially sent a randomly selected set of 238 tracing cases to both this vendor and a vendor specializing in database searches. The database search vendor conducted electronic searches on a variety of databases to which they have access. These databases include telephone directory listings, credit records, death records, and other similar types of databases.

The results of this test were roughly equivalent in terms of the number of cases found by each of the two vendors using their respective procedures. The telephone number look-up vendor found 39% of the cases submitted while the database search vendor found 38%. However, the cost of the telephone number look-up search was only a fraction of the cost of the database search. Thus, we decided to continue using the former for this step in the tracing process.

In-Depth Searches. The in-depth searches consisted of the electronic database search step just described. In addition, however, the vendor's staff did more intensive tracing that started with the last known address of the selected physician. Physicians were then traced by contacting

neighbors, prior employers, and other potential contacts to establish a current telephone number and/or address.

These in-depth search procedures were used in one study site with an exceptionally high number of tracing cases. They were also used to trace a random sample of unlocatable physicians. This effort is described in Chapter 5 as part of the response rate discussion.

Internet Searches. Four states maintain physician registries accessible by the public on the Internet. The Internet site is called AIM and is prepared by the Association of State Medical Board Executive Directors. Find rates varied by state, depending on the type and timeliness of the data provided.

4.7 Refusal Conversion

Physicians are notoriously difficult to interview. They frequently refuse either in person or through their office staff. If the refusal was "soft," Gallup interviewers merely held onto the case for awhile, then tried again. For example, a soft refusal might be a physician telling the interviewer he is too busy to do the interview or a receptionist saying the doctor doesn't do surveys. In the former case, the interviewer would emphasize his/her flexibility with regard to scheduling an interview at a convenient time. (One interviewer actually completed an interview at 3 a.m. because that was the only time an Emergency Room physician had time to do it.) However, if the physician persisted in saying he/she did not have time, the interviewer would put the case aside for a few weeks or longer, then try again. Frequently, the later attempt would catch the physician at a better time and the interview could be completed. In the example of the receptionist who screens out the interviewer's call by saying the doctor doesn't do surveys, the interviewer would try to call at a time of day when the receptionist was not there. Sometimes a different office worker would put the call through to the doctor and sometimes the doctor would answer the telephone him/herself and could be persuaded to complete the interview.

Soft refusals of the types just described were usually coded by the interviewers as callbacks, not refusals. Sometimes, as in the example of the case to be held for awhile, the case would be coded as a "First Refusal." In both instances, however, the ownership of the case stayed with the original interviewer who would follow-up on the case at a later time. None of these types of soft refusals are included in the following discussion of refusal conversion efforts.

There are two types of refusals that are included in refusal conversion efforts: hard refusals and second refusals. A hard refusal is coded when the physician or office worker becomes so hostile that the interviewer considers it necessary for future attempts to be handled by a refusal conversion specialist. Second refusals are cases where two soft refusals have been received. Again, these cases are best referred to a refusal conversion specialist.

In this study, there were 5,635 hard or second refusals which is 23.9% of all released cases. Gallup's refusal conversion team converted 1,199 (21.3%) of these refusals to completed interviews. Another 281 were re-categorized from refusal cases to ineligible, unavailable during the study period, too ill to participate, or some other final status code. This reduced the final refusals to 4,155 or 17.6% of all released cases. The following paragraphs describe Gallup's refusal conversion efforts.

First, it should be understood that Gallup's executive interviewing team is a group of highly experienced interviewers who are good at avoiding refusals and converting soft refusals into completed cases. Therefore, cases that become hard or second refusals have already been worked very hard. On studies that require less experienced interviewers, it is often the case that a high proportion of refusals can be easily converted by a more experienced interviewer. This is not the case in this study. All of the refusals assigned to the refusal conversion team were very difficult cases that, in all likelihood, could not be converted easily.

The procedure Gallup followed in working the refusal cases was as follows. First, we allowed refusals to age as long as possible. Although the field period began mid-August, we did not assign the first refusal conversion cases until December when about three months had passed since the first refusals occurred. In our experience, physicians are so busy that they sometimes forget about the study and the earlier interactions which means refusal conversion attempts can be approached as though it is a new study. Rather than reminding the physicians of their prior refusals, therefore, we merely sent a second copy of the introductory letter and the refusal conversion specialist began calling the doctors as they would new cases though armed with complete information about the prior difficulties. At the time of the refusal, the original interviewer entered notes describing the interaction(s) that were later used by the refusal conversion specialists in formulating their strategic approach to the conversion attempt.

The Gallup refusal conversion team consisted of 11 of the most experienced and effective physician interviewers. Some interviewers have a special interest in refusal conversions. These interviewers tend to have a knack for getting past gatekeepers and for effectively addressing physicians' concerns to "turn around" these difficult refusal cases. They derive particular satisfaction out of converting refusals even though they are unsuccessful in doing so in the majority of cases. Other excellent interviewers prefer not to do refusal conversion work. They find it demoralizing even if they are quite successful at doing it. For this reason, we only assign refusal conversion cases to interviewers who want to work on them.

4.8 Respondent Incentives

The initial incentive plan was to send a check for $25 to each physician after he/she completed an interview. Our advance letter offered the $25 honorarium and explained that it would be paid upon interview completion.

Although prior research has not found that sending incentive money in advance of interview completion improves response rates, there is some evidence suggesting advance payment may speed up the response process[5]. That is, ultimately, we might expect the same total number of respondents, but with advance incentives we might be able to reduce the length of the field period. For this reason, it was decided about halfway through the field period to send advance incentive checks to the remainder of the open cases in the hope that a faster response time could be achieved.

Advance incentive checks were sent to 8,485 open cases about three-quarters of the way through the field period. Anecdotally, interviewers felt that the advance checks did help them to gain access to physicians. However, we do not have empirical data that bears on the question of whether the advance incentive procedure produced faster response times.

Physicians who received incentive checks after completing the interview were more likely to cash the checks than physicians who received incentive checks in advance were. Among those who received checks after the interviews, 94% cashed the incentive checks. Of those who received checks in advance, just 88% cashed the checks. The combined percentage of physicians who cashed their checks was 90%.

---

[5] Shettle, Carolyn, and Geraldine Mooney, "Evaluation of Using Monetary Incentives in Government Studies," Washington Statistical Society, November 19, 1996.

4.9 Physician Recruiters

Six months after the study began, we enlisted the help of a physician to contact refusal cases and hard to reach respondents. Our aim was to evaluate the efficacy of employing physicians to help recruit respondents.

The physician recruiter was assigned 99 cases, 53 hard refusal cases and 46 cases that had been attempted more than ten times without reaching the respondent. The physician recruiter obtained verbal agreement to complete the interview from 11 of the 53 hard refusal cases (21%) and from 17 of the 46 hard to reach cases (37%).

Once the physician recruiter obtained verbal agreement, the case was sent back to the original interviewer to contact the respondent and complete the interviews. Interviewers were able to complete three interviews with the hard refusal cases (5.7% of 53) and eight interviews with the hard to reach cases (17.4% of 46).

Our conclusions from this limited experiment were that the physician recruiter might be helpful in obtaining interviews with the hardest cases. However, this method of recruiting respondents is costly. It would be imperative to conduct the follow-up calls immediately after the physician recruiter obtains the respondent's agreement to participate instead of some days later as in the procedure we used.

5. Production Statistics and Quality Control

Table 2 shows the final disposition of all unduplicated released cases for the Community

Tracking Study Physician Survey.

Table 2
Final Disposition of Unduplicated Released Cases

Status	Label	Frequency	Percent	Eligibility
1	Complete	12,385	53.6	Eligible
12	Language Barrier	6	0	Unknown
13	Hearing Barrier	4	0	Unknown
15	Respondent too ill to participate	33	0.1	Unknown
17	Final Refusal	4,155	18.0	Unknown
18	End of Study	1,310	5.7	Unknown
19	Ineligible (Federal employee, resident or fellow, excluded specialty)	400	1.7	Ineligible
20	Final Tracing/Unlocatable	2,751	11.9	Unknown
22	Deceased	102	0.4	Ineligible
24	Respondent unavailable during field perieod	170	0.7	Unknown
25	Final, Other reason	9	0.0	Unknown
26	Corporate Refusal	11	0.0	Unknown
27	Retired	913	4.0	Ineligible
28	>20 hrs/week direct patient care	847	3.7	Ineligible
	TOTAL	**23,096***		

*This total excludes the 525 "No Contact" cases which were included in the weight calculations but were excluded from the fielded sample, a practice consistent with the AMA's Socioeconomic Characteristics of Medical Practices Study.

The table above shows the final status codes assigned with a description of each code. The

number of cases assigned each of the final status codes appears in the "Frequency" column

followed by the percent of sample. The last column of the table shows how each of the status

codes maps to respondent eligibility. This mapping holds true with the following exception. In a

few cases, respondents completed the screener and were known to be eligible, but did not

complete the interview. For example, a respondent might have completed the screener, but then refused to do the interview. In these cases, the particular respondent record is coded as known eligible even though his/her status code would map into an unknown eligibility. There were only 72 such cases.

An important aspect of response rate calculation is determining the presumed eligibility rate among cases with unknown eligibility. In general, it is presumed that respondents whose eligibility is unknown are eligible at the same rate as respondents for whom eligibility has been determined (known eligibles + known ineligibles). In this study, we had reason to question that the eligibility rate among physicians who could not be located was the same as the eligibility rate among uncooperative physicians we did locate. For this reason, we conducted an in-depth tracing effort to determine the eligibility rate among unlocatable physicians. This effort is described in the next section.

5.1   In-Depth Tracing of Unlocatables

After the normal tracing procedures, internet searches, and special searches were completed, a methodological experiment was conducted to determine, to the extent possible, the eligibility rate among the 2,751 unlocatable physicians. A random sample of 400 unlocatable physicians was selected. In-depth searches were conducted to locate these 400 cases. The in-depth searches included extensive searching of databases (as described previously in Section 4.6). It also included contacting neighbors and employers for current addresses and telephone numbers or for forwarding information if the physician had moved. Information obtained was followed up to determine a current telephone number and/or address.

If a possible telephone number was identified for a physician, Gallup interviewers conducted confirmation calls to verify that the located physician was the correct sampled physician. In

addition to verifying the physician's full name, interviewers verified the physician's primary specialty and birth year, and asked the screener questions to determine eligibility.

In some cases, a current address was found, but no telephone number. For example, sometimes it was possible to find a current residential address in credit records, but the telephone number was unpublished. In these cases, Gallup sent a letter explaining the purpose of the study and asking the physician to complete and return a one-page questionnaire containing the screening questions. The letter explained that it was important for us to gather the screening information even if the physician preferred not to participate in the interview. The physician had the option of providing his/her telephone number and could indicate whether he/she preferred not to participate.

The results of these in-depth searches are listed in Table 3 below:

Table 3
Results of In-Depth Searches

Outcome	Frequency
Eligible	**107**
Not Eligible	**63**
Retired	17
>20 Hours/wk. patient care	18
Moved out of country	6
Resident/ Fellow	14
Federal employee	5
Deceased	3
Eligibility Unknown	**230**
Not located	159
Sent letter/No reply	66
Refused screening questions	5
TOTAL	**400**

These results suggest that a conservative estimate of the eligibility rate among unlocatable physicians is 62.9% (107 of 170 located physicians). This is a conservative estimate because we were unable to find any record of 159 of the sample of 400 physicians (39.8%) even after extraordinary tracing efforts. The following are typical search reports:

> Automated files provided no new information. Directory assistance had no listing for the subject at the address provided. Neighbors were not familiar with the subject. A surname search in the U.S. found no one with the same name.

> Consumer files had no record of the subject. Directory assistance had no listing for the subject at the last known address. Neighbors had no information. A surname search in the state showed 6 persons with the same name. They were contacted, but did not know the subject.

At a minimum, these search reports suggest that the 159 unlocated physicians are not currently practicing medicine. In many cases, it is likely that they are no longer living in this country. If we assumed that the 159 unlocated physicians were not eligible, the eligibility rate among unlocatable physicians would decrease to 32.5% (107/107+63+159). The denominator used in this calculation is 329 (400 physicians minus 71, those who did not reply to our letter or who refused to answer the screening questions).

Following the logic of the preceding paragraphs, we can reasonably hypothesize that the true eligibility rate among unlocatable physicians is somewhere in the range of 32.5% to 62.9% eligible. In subsequent response rate calculations, we have taken the most conservative approach by using 62.9%, the highest level in this range, as the eligibility rate among unlocatable physicians.

5.2 Response Rate Calculations

5.2.1 Overall Response Rate.

The overall response rate for the Community Tracking Study Physician Survey-including both the site and Supplemental samples-was 65.4%.[6] This section describes the calculation of the response rate.

The response rate (R) is the proportion of eligible cases that complete an interview:

$$R = C/E$$

where C is the number of sample cases who completed an interview and E is the number who were eligible.

Determining the exact number of physicians who were eligible for the study is complicated by the large number of sample physicians whose eligibility was never established. Based on the results of the methodological study described in Section 5.1, we estimated the eligibility rate among the unlocatable physicians to be 62.9%. In addition, we estimated the eligibility rate among the physicians who were located but who did not complete a screener to be 84.6%. This latter figure was the proportion of eligible physicians among all physicians who were successfully screened.

Table 4 shows the data on which the response rate calculations were based.

---

[6] Approximately 2% of physicians will not permit AMA to give out their names, addresses, or telephone numbers for any purpose Although these physicians are included in the calculation of weights for this survey, they are not included in the response rate calculations since they had no opportunity to participate in the study. Response rates for physician studies conducted by AMA and other researchers follow this same procedure.

Table 4
Response Rate Calculations

Outcome	All Physicians	Primary Care Physicians`	Non-Primary Care Physicians
Completed Cases (a)	12,385	7,634	4,751
Known Eligibles (b)	12,457	7,682	4,775
Estimated Eligible Unlocatable Physician (c)	1,729	1,200	529
Estimated Eligible Locatable Physician (b)	4,761	3,049	1,712
Total Eligibles (e=b+c+d)	18,947	11,931	7,016
Response Rate (a/e)	65.4%	64.0%	67.7%

Overall, 12,385 sample physicians completed an interview. Besides the physicians who completed the main interview, 72 nonresponding physicians were known to be eligible based on their responses to the screener. Two of these nonrespondents moved prior to the main interview and became unlocatable. The number of eligible cases among the remaining physicians had to be estimated. There were a total of 2,749 unlocatable physicians whose eligibility was not determined. Using the results of the in-depth search of unlocatables (Section 5.1), 1,729 of these (62.9%) were estimated to be eligible for the study. Another 5,628 physicians were located but did not complete the screening interview. 4,761 of these (84.6%) were estimated to be eligible, based on the proportion of eligible physicians among all who were successfully screened. Altogether, then, our estimate of the total number of eligible physicians was 18,947 (that is, 12,457 plus 1,729 plus 4,761).

The overall response rate-65.4%-was the number of completes (12,385) over the estimated total number of eligibles (18,947). Parallel computations yielded estimated response rates of 64.0% among the primary care physicians and 67.7% among the non-primary care physicians.

5.2.2 Item Non-Response.

In a CATI interview, it is not possible for respondents to skip items. However, they may refuse to answer a question or they may indicate that they do not know the answer to a question. We have reviewed the frequency of these two types of item non-response. Both types of problem were relatively rare. Over 99% of respondents were missing fewer than five items.

Item refusals were generally rare. Few of the individual survey items had high rates of missing data due to refusals. One exception was, as expected, the income question (H10) which 8.9% of physicians (n=1,107) initially refused to answer. Of these, however, more than half provided an income category in response to the follow-up question (H10A). Thus, combining across the two questions, just 3.5% of respondents failed to provide any income information.

Throughout most of the interview, "don't know" responses were also rare. Higher levels of "don't know" occurred in Section G (Practice Revenue). The range for "don't know" responses for the main Section G questions (G1a, G1b, G3, G6, G7, G8, G9, G11) was 7.1% (G3, percentage of practice revenue which is capitated or prepaid) to 16.5% (G9, percentage of practice revenue from largest contract).

The only other question in the survey that elicited a notable percentage of "don't know" responses was question B6, the number of hours spent providing charity care in the past month; 6.8% answered "don't know" to this question. For all of the remaining questions in the survey (excluding follow-up questions), the rate of "don't know" responses was 2.6% or less.

5.3 Data Preparation

Most of the data coding and cleaning was accomplished by the CATI system. As the interviewers entered response option codes selected by the respondents, these numbers were written to a data file. The CATI system was programmed to conduct range and consistency

checks and to prompt the interviewer when an impossible or unlikely response was entered. The interviewer could then correct the data entry or ask the respondent to clarify his/her answer.

Range Checks. The ranges of most closed-ended items in a CATI survey are determined by codes for the available responses. For example, a "Yes/No" variable offers the following codes:

    1   =   Yes
    2 =   No
    8 =   Don't know
    9 =   Refused

If the interviewer mistakenly attempts to enter a code of "3," the CATI system will notify the interviewer that this is an unacceptable code. The interviewer can then reenter the correct code.

Some items such as dates, number of hours worked, or percentages of revenue, do not have a set of preassigned response codes. Ranges are bounded by what is possible. For example, question B1 asks the respondent how many weeks he/she practiced medicine during 1995. Since there are 52 weeks in a year, the acceptable range for responses was 00 to 52. Higher numbers were not accepted by the system.

Consistency Checks. Consistency or logic checks examine the relationships between two or more variables to be sure that the responses do not conflict with one another. A few logic checks were contained in the CATI program. For example, in Section B, question #B2 asks the physician how many hours he/she spent in all medically related activities last week. Then, question #B3 asks how many hours he/she spent in direct patient care last week. If the responses to these two questions are equal, a check question is asked to be sure that all of the physician's time was spent in direct patient care. Alternatively, if the physician indicated he/she spent more hours in direct patient care, than in all medically related activities (a logical impossibility), the physician is prompted to revise one or both of the answers to questions B2 and B3.

Section G of the questionnaire also contains several consistency checks. Check questions appear if the combined practice revenue from Medicare and Medicaid is greater than 100%; if the revenue from all managed care contracts is less than the amount received on a capitated basis; if all of the practice's managed care revenue is paid on a prepaid basis; if the percentage of revenue from the practice's largest managed care contract is greater than the total revenue from all managed care contracts; if the practice has more than one managed care contract, but the revenue from the largest managed care contract equals the total revenue from all managed care contracts; and if the physician says that his/her practice has more than 20 managed care contracts.

In a survey as complex as the Physician survey, the potential number of consistency and logic check questions is very large. It was decided during the questionnaire development phase that the number of such questions to be programmed into the CATI instrument was to be limited to only the most important questions. The reasons for this decision were that the time required to program and test such consistency check questions is considerable and that very few respondents were expected to be affected. This reasoning is substantiated by the item frequencies. Of 12,385 completed interviews, only 7 responses were corrected in the Section B check series. However, the Section G questions about practice revenue generated a bit more confusion. Inconsistent responses to several questions were corrected for over a hundred respondents.

5.3.1 *Data cleaning.*

Although most data cleaning is done on-line with a CATI survey, there are a few data cleaning steps to complete when the survey comes out of the field. Frequencies are examined and cross-tabulations are run to check for additional consistency checks that were not built into the survey. On the basis of these tabulations, data may be changed or flagged for further checking. Occasionally, a check step may have been overlooked during CATI development which requires cleaning of the data once the survey is out of the field. For example, on this survey, we failed to

program the CATI to reject respondents currently practicing medicine in Alaska or Hawaii. Thus, after the survey was completed, four completed cases were dropped because the respondents had moved from the location where they were originally sampled to practice in Alaska or Hawaii. Similarly, one physician whose specialty was among those excluded from the study was interviewed because of an oversight in the CATI program.

5.3.2   Coding.

The amount of post-interview coding for this survey was very limited. Four questions in Section C permitted entry of "Other, Specify" responses (Questions C2, C3b, C6, and C6a). These open-ended responses were then examined after the survey came out of the field to determine whether the responses given by the respondent actually fit into the categories provided in the question. If not, no change was made. If the response did fit an existing category, the "Other, Specify" response was recoded to the correct response category. A few response categories were added to permit coding of most of the "Other, Specify" responses.

The nature of the sample for this study made it very important to consider the location of each respondent's practice at the time of interview in comparison with the location at the time of sampling. Physicians in the site sample were sampled as part of the population of a particular site. Each site was defined as containing a particular set of FIPS codes. During the interview, every respondent was asked to confirm the location (county and state) of his/her primary practice. Respondents whose practices were not located in the county and state shown in the sample record, were asked to provide their current county and state.

6. Weighting of Data

6.1 Overview

The total number of cases (respondents and non-respondents) in the sample used for the calculation of weights was 23,621, consisting of (i) 23,096 cases released for interviewing and (ii) 525 "do not contact" cases. Because the site sample and supplement sample were independently selected, some physicians (265 cases) were sampled in both the site and the supplemental samples. Only one of these selections are represented in the released count of 23,096 cases. For weighting purposes, these cases are accounted for in the site sample and the supplemental sample as appropriate to the particular weighting situation.

The number of cases belonging to selected sampling subgroups is summarized in Table 5 below.

Table 5
Number of Cases in Selected Subgroups

		Number of cases sampled
Sample Component		
Site Sample		21,616*
Supplemental Sample		2,270*
	Total (includes duplicates)	23,886*
Primary Care/Non Primary Care Specialty		
Primary Care		15,135
Non-Primary Care		8,486
	Total (excludes Duplicates)	23,621

* Includes 265 duplicate cases (i.e., physicians selected for both site and supplemental samples).

The survey design includes two statistically independent but overlapping nationally representative samples and permits the development of a variety of weights for analysis at the national level in addition to site-level analysis. Six different sets of weights, outlined in Table 6 below, were calculated for the physician survey. After describing the general approach used to construct the weights, this section provides the specific details on the construction of each type of weight.

Table 6
Types of Weights Calculated for Physician Survey
(N.B. Situation 3 Omitted; Not Calculated)

Name	Physicians Included:	Type of Estimate	Completed Interviews
Situation 1	Drawn from site sample and, when surveyed, Located in one of the 60 sites	Site	10,881
Situation 2	Drawn from site sample	National	11,310
Situation 4 (WTPHY3)*	Drawn from supplemental sample	National	1,218
Situation 5 (WTPHY1)*	Drawn from site and supplemental samples and, when surveyed, located in one of the 60 sites.	Site	11,474
Situation 6 (WTPHY2)*	Drawn from site sample and, when surveyed, located in one of the 60 sites.	National	10,881
PHNATLWT (WTPHY4)**	Drawn from site and supplemental samples. Includes all physicians surveyed.	National	12,528

*Weight variable names used in the Physician Survey Restricted Use File
**Weight variable name used in the Physican Survey Public Use and
   Restricted Use Files.

6.2 General Approach

Three components of the sampling weights were computed: (i) the probability component weight or the base weight, (ii) the nonresponse adjustment, and (iii) the post-stratification adjustment. The sampling weights were also trimmed to reduce the variance inflation effects of extreme weights.

The probability component weight assigned to a physician equals the reciprocal of the probability of inclusion of that physician in the sample. The inclusion probability was determined based on the sampling stratum in which the physician was assigned. In the site sample, four sampling strata were defined within each site by crossing the two specialty groups (primary care physicians (PCP) and non-primary care physicians (NPCP)) with the two sample sources (AMA & AOA). In the supplemental sample, two sample source strata (AMA & AOA) were defined within each of the twenty national strata formed by crossing ten geographic locations (groups of states) and two specialty groups (PCP & NPCP). The population counts at the time of sampling were used for calculation of inclusion probabilities. The specialty code that was available in the AMA/AOA source file for each physician was used to identify each physician as a primary care (PCP) or non-primary care (NPCP) physician.

In some cases, the location of the physician office address when surveyed may have been different from the address listed on the AMA Master File as the primary mailing address. In constructing the sampling weights, we took into consideration both the site when sampled (as reported in the frame) and the site reported by the physician as the location of the physician's primary practice. (In situations involving only the supplemental sample, the stratum where sampled and the stratum reported in the survey were taken into account.) With reference to a particular site or stratum (for example, Site A), the sampled physicians were grouped into the following categories: (i) non-movers: physicians for whom the site when sampled (Site A) was

the same as the reported site, (ii) in-movers: physicians who were not sampled in Site A, but reported to be in Site A in the survey, and (iii) out-movers: physicians who were sampled in Site A and then found to be in a different site or location at the time of interview.

The survey question on "location of practice" was asked after asking the screening questions, i.e., after establishing the eligibility of the respondent. Hence, for the survey only "known eligible" physicians were classified into these three categories (non-mover, in-mover, and out-mover). For any physician in these categories, we knew the site when sampled and the site reported. For any physician who was not asked the "location of practice" question (the interview was terminated before that stage either because the respondent was found ineligible or the respondent refused to continue), information on actual practice location was not reported in the survey. In these cases, the physician was treated as a non-mover.

To calculate the nonresponse adjustment weight, nonresponse adjustment cells were defined to include responding and non-responding physicians based on the reported site (or national stratum in situations involving the supplemental sample only), physician's specialty, and age. The two specialty groups were PCP and NPCP, and the two age groups were "less than 50 years" and "greater than or equal to 50 years." The goal was to keep the minimum number of physicians in nonresponse cells to around 20. Within each nonresponse adjustment cell, the sampled physicians were grouped into three groups: (i) known eligibles; (ii) known ineligibles, and (iii) those with unknown eligibility. Out-movers were considered ineligible in the site where they were sampled and eligible in the main practice site reported during the interview. The 525 "do not contact" cases were put in the "eligibility unknown" category.

The weighted response rate within each nonresponse cell was computed as follows:

$$R = C/E \qquad\qquad (1)$$

51

where C is the weighted count of sample cases who completed an interview and E is the weighted count of cases who were known to be eligible plus an estimate of the number of eligibles among the cases where eligibility was unknown. The weighted counts in both the numerator and the denominator used the probability weight component. All completed interviews from the non-mover and in-mover categories were included in the numerator. All known eligibles from non-mover, in-mover and other nonrespondent categories were counted under "known eligibles." The estimated number of eligible physicians among those for whom eligibility was unknown was calculated slightly differently for weighting purposes than the calculations for response rates described in Section 5.2. For weighting purposes, the eligibility rate among all cases for whom eligibility was unknown was estimated, within each nonresponse adjustment cell, by the eligibility rate among cases for whom the eligibility was known.

For the purpose of calculating the weighted response rate (to adjust for nonresponse), we could have used (i) the specialty based on the frame variable for all physicians or (ii) the self-reported specialty for physicians who completed the survey and the frame variable for nonrespondents. Using the second approach would provide only a partial correction because it was not known whether physicians who did not cooperate in the study would have reported a different specialty or not. Further, we realized that the second approach would be quite complicated as compared to the first. Adjusting the numerator (the number of completes) of the response rate would be straightforward. However, estimating the number of eligibles in the denominator would be complicated and would not correct the entire sample. Therefore, we decided to adopt the first approach and use the specialty variable on the frame for all physicians for the purpose of the non-response adjustment.

The final component of the sampling weight involved trimming the response-adjusted sampling weights and computing poststratification adjustments. Details of the calculation of sampling weights are described below for each situation separately.

6.3 Sampling Situations

6.3.1   Situation 1: Site analysis based on site sample.

In this case, a set of weights was developed for all physicians who were initially drawn from the site sample and whose practice location was in one of the 60 sites. For analysis of a specific site (say Site A), all completed interviews from the non-mover and in-mover categories were included in the Site A sample. Out-movers were considered ineligible and hence were excluded from the Site A analysis. They were, however, included in the calculation of weighted response rates (for nonresponse adjustment calculations) in Site A.

The probability weight component assigned to a physician sampled from one of the four sampling strata (PCP/NPCP x AMA/AOA) within a site was calculated as N/n where "n" and "N" were respectively the sample and population size of that sampling stratum. In the case of in-movers, this probability was multiplied by the probability of inclusion of the original site (where the in-mover was originally sampled) in the sample of sites. This was necessary because the probability of inclusion of an in-mover in the site sample is equal to the product of (i) the unconditional probability of selection in the site of origin in the first stage sample of sites and (ii) the conditional probability of selection of the in-mover in the site sample given that the corresponding site is already chosen in the first stage sample of sites.

For calculation of nonresponse adjustment weights, four nonresponse adjustment cells were formed within each site by crossing two specialty groups and two age groups. The minimum cell size was 17 after merging two pairs of cells in one site. No other merging of nonresponse cells

53

was necessary in any other site. Formula (1) (Section 6.2) was used to compute the weighted response rate. The nonresponse adjustment component of the weight was calculated as the inverse of the response rate. The response-adjusted weight assigned to a case with a completed interview was the product of the probability weight and the nonresponse adjustment component.

Upon inspection of the distribution of the weights and weighted site counts, we found that "movers" (i.e., physicians whose reported practice location differed from the location where they were sampled) resulted in some anomalies in the population estimates of eligible physicians in some sites. In addition, the process used to account for physicians "moving" into a site from another site produced sampling weights that were sometimes much larger than the sampling weights for the "non-mover" physicians in the site. Because most of the anomalies in the site population estimates seemed to have been caused by these excessively large weights, an algorithm was used to trim the large sampling weights and achieve a more consistent estimate of the site population of eligible physicians. The weight-trimming algorithm compared each weight to the square root of the average value of the squared weight. This algorithm has been referred to as the "NAEP" procedure[7]. After trimming some of the large weights, the sum of the weights was used as the population estimate for the site. The details of the algorithm used for trimming weights in Situation 1 are given below.

Two specialty subgroups (PCP and NPCP) were created in each of the 60 sites, resulting in 120 subgroups. Sampling weights were trimmed using the following steps:

(1)   The NAEP criterion value (NAEP1) was calculated for each subgroup using the

following formula:

$$\text{NAEP1} = \text{SQRT} [c* (\text{Sum of squared weights})/ n] \tag{2}$$

[7] Potter, F.J. "A Study of Procedures to Identify and Trim Extreme Sampling Weights." *Proceedings of the American Statistical Association, Section on Survey Research Methods,* 1990, pp. 225-230.

where c=10 and n is the size of the subgroup. Any weight greater than NAEP1 was trimmed by truncating the weight at the value of NAEP1. The untrimmed weights were not changed.

(2) The NAEP criterion value (NAEP2) was recalculated for each subgroup based on the trimmed weights obtained after carrying out step (1). Any weight exceeding NAEP2 was truncated at NAEP2. The untrimmed weights were not changed.

(3) The NAEP criterion value (NAEP3) was calculated based on the trimmed weights obtained after carrying out step (2). Any weight exceeding NAEP3 was truncated at NAEP3. The untrimmed weights were not changed. The final weights were those obtained after carrying out step (3).

The trimmed weights were then post-stratified so that the weighted estimates of the number of PCP and NPCP physicians within each site in Situation 1 matched the corresponding estimates in Situation 5 (site estimates using site sample and supplemental sample). It was determined that the population estimates in Situation 5 were our best estimates because Situation 5 contained additional sample units (i.e. more information) compared to the Situation 1 sample. The weighted estimates of the population total of PCP and NPCP groups within each site in Situation 1 were ratio-adjusted to the Situation 5 best estimates.

The sampling weights at this stage were then subjected to a second round of weight trimming to address the potential of extreme weights to inflate the sampling variance of survey estimates. Within each site, two trimming classes (PCP and NPCP) were again used. The NAEP procedure was used again with an assessment of the impact of trimming on the sampling variance. This step identified weights to be trimmed and distributed the trimmed excess among the weights that were not trimmed. The statistical measure of the impact of the trimming was based on the design effect attributable to the variation among the sampling weights. The design effect attributable to weighting is a measure of the potential loss in precision caused by the variation in the sampling weights relative to a sample of the same size with equal weights. Sampling weights were trimmed to reduce the design effect, and yet minimize the risk of introducing bias into the

55

sample estimates, that is the extent of trimming was limited to ensure minimal effect on survey estimates.

6.3.2   *Situation 2: National analysis based on site sample*

In this instance, the entire site sample, including out-movers, was used to develop weights for national estimates. The site sample is a two-stage probability sample drawn from the national frame (of the population of all physicians). In the first stage, a probability sample of 60 PSUs (sites) was chosen from the frame of all sites. In the second stage, random samples were drawn independently in each of the 60 sites chosen in the first stage sample. Hence, the site sample is a nationally representative probability sample of all physicians and can be used to generate statistically valid weighted estimates at the national level and the precision of these estimates.

The sample of 60 PSUs consisted of 48 medium and large MSA PSUs, three small MSA PSUs and nine other non-MSA PSUs. Among the 48 medium and large metro PSUs, 12 were selected at random with equal probability to be high-intensity sites, and the rest--the other 36 medium and large MSA PSUs plus the three small MSA and nine non-MSA PSUs--were designated low intensity sites. The calculation of the inclusion probability (Pi) for any sampled physician took into account the way high intensity sites were chosen. The probability of selection of physician i from any one of the four sampling strata within a site was calculated using the following formula:

$$P_i = P(PSU)*P(i|PSU) = P(PSU)*[P(HI)(n_{HI} / N_s) + (1-P(HI))(n_{LO}/N_s)] \tag{3}$$

where Ns was the population size (of the sampling stratum), $P(HI) = 12/48=1/4$ for the 48 large metro PSUs and zero for the rest, n_{HI} (n_{LO}) is the sample size that would have been allocated to a site if it was chosen as a high (low) intensity site. The use of formula (3) in each of the four sampling strata (PCP/NPCP X AMA/AOA) within each of the 48 large MSA PSUs required the

estimation of the sample size that would have been released under our original sample allocation plan treating each site first as high-intensity and then as a low-intensity site.

Based on the original sample design and allocation plan, we assumed target effective sample sizes for primary care physicians to be 400 and 100 for high-intensity and low-intensity sites, respectively. For non-primary care physicians, the corresponding target numbers were 191 and 51 physicians. These numbers were not fixed for each site in the original allocation plan but were determined based on other constraints, such as the precision for estimates for all physicians in the site.

We then computed the sample size required for each high intensity site as if it was a low intensity site and vice versa. We made this computation using the ratio of the numbers released for each site (within each stratum) under the assumption of high and low intensity site. This ratio was then used to estimate the unknown sample size for a high (or low) intensity site given the known sample size for a low (or high) intensity site. The estimated sample sizes were used in the calculation of selection probabilities based on formula (3).

Nonresponse adjustment components were calculated following an approach similar to that used in Situation 1. In Situation 1 physicians who reported that they were not in any of the 60 sites were excluded from the analysis. For Situation 2, however, physicians who reported themselves to be outside any of the 60 sites were still part of the sample representing the national population and hence were included in national analysis. All physicians not found to be in one of the 60 sites at the time of interview were considered to be in a separate site in order to be able to use the nonresponse weighting procedure (based on reported site) used in Situation 1. Calculation of weighted response rate and the nonresponse adjustment component was similar to Situation 1.

The sampling weight at this stage was the product of the probability and the nonresponse weight components. These weights were then post-stratified so that the weighted national estimates of the total number of PCP and NPCP physicians obtained in Situation 2 matched the corresponding estimates in Situation 4 (based on the supplemental sample only). The Situation 4 estimates were considered to be our "best national estimates" because supplemental sample was a simple stratified sample that was not affected by movers or multiple stages of selection. The Situation 2 estimates were ratio adjusted to these "best estimates."

Using the post-stratified weights, we then trimmed the weights (using NAEP criteria) within classes formed as follows. For large metro areas, 16 trimming classes were formed by crossing 4 geographic regions with 2 levels of physician type (AMA/AOA) and 2 levels of specialty code (PCP/NPCP). For small metro areas, 2 trimming classes (PCP/NPCP) were used. For non-metro areas, 8 trimming classes were formed by crossing 4 geographic regions with 2 levels of specialty type (PCP/NPCP). After a weight was trimmed, the excess weight was redistributed among the untrimmed weights in the class.

6.3.3   Situation 3: Not used

6.3.4   Situation 4: National analysis based on supplemental sample.

In this situation, the supplemental sample is used by itself as a basis for national estimates. The supplemental sampling frame consisted of all physicians stratified into 20 strata based in 10 geographic regions (groups of states) and 2 specialty types (PCP/NPCP) within each region. Simple random samples were drawn from each of these 20 strata independently from the AMA and AOA lists of physicians. The supplemental sample, therefore, was a simple stratified random sample drawn from the national frame and hence was a representative sample of all physicians.

The probability component weight assigned to a physician sampled from one of the two sampling strata (AMA or AOA) within any one of the twenty national strata (10 state groups by two specialty groups) was calculated as N/n where "n" and "N" were respectively the sample and population size of that sampling stratum.

For calculation of nonresponse adjustment weights, four nonresponse adjustment cells were formed within each of the twenty national strata by crossing two specialty and two age groups. Formula (1) (Section 6.2) was used to compute the weighted response rate and the nonresponse adjustment component of the weight was calculated as the inverse of the weighted response rate. The final weight assigned to a case (with a complete interview) was the product of the probability and the nonresponse component weights. No post-stratification adjustments were done to weights in Situation 4.

The set of final weights were then trimmed using four trimming classes obtained by crossing two levels of physician type (AMA/AOA) with two levels of specialty (PCP/NPCP). After a weight was trimmed, the excess weight was redistributed among the untrimmed weights in the class.

6.3.5 *Situation 5: Site analysis based on combined (site and supplemental) sample in 60 sites.*

In this situation, weights suitable for site analysis were developed based on all those physicians in the site sample and all those in the supplemental sample whose practice location when surveyed was in one of the 60 sites. Physicians whose practice location when surveyed was not in one of the 60 sites were excluded.

In order to calculate the probability component of the weight, we derived from the combined sample of physicians the number sampled (n) in each of the four sampling strata within each site. Physicians who were sampled twice (in both the site and supplemental samples) were duplicated. The probability weight was calculated as N/n where N was the corresponding stratum population

size. For in-movers, this weight was also multiplied by the first stage inclusion probability of the site where the in-mover was originally sampled. The calculation of the nonresponse weight component was similar to Situation 1. There were no post-stratification adjustments to the sampling weights in Situation 5. The final weights were then trimmed using the trimming procedure described in Situation 1.

6.3.6 *Situation 6: National analysis based on site sample excluding physicians with practice Locations outside the 60 sites.*

The sampling and the weighting procedures were the same as those used in Situation 2, except that physicians whose practice location was outside the 60 sites at the time of interview were excluded from analysis in Situation 6.

6.3.7 *PHNATLWT (WTPHY4): National analysis based on entire combined site and supplemental samples.*

The Community Tracking Study includes two sample components: a national multistage sample using 60 sites and a national supplemental sample. Point and variance estimates based on the combination of these two samples can be constructed using estimates computed from the site and supplemental samples separately and then combining the estimates to form national estimates. (See Section 6.3.7.1) This strategy provides the most accurate point estimates in the sense of minimizing the estimates of the sampling variance. However, it also results in some discrepancies in the analyses (for example, the sum of percentages does not always add to 100 percent) and involves additional processing time. Furthermore, this strategy is difficult to implement for regression-type analyses.

In view of these difficulties, a strategy was explored to combine the two sample components by adjusting the weight for each sample so that the sum of the weights across the two samples equals the population total. The purpose of the effort was to find one or more values of a scaling factor (called lambda) that could be used to combine the weights from each sample component and achieve the best estimates with nearly minimal sampling variances for these estimates as

well as reducing the computer processing. Conceptually, any value of lambda would result in unbiased estimates, but the "best" point estimate is associated with the value of lambda that achieved the minimum variance. The effort, therefore, was directed at the identifying a value of lambda that achieved the smallest variance estimates across various subpopulations and analysis variables.

The estimation of the scaling factor used variance estimates computed for each component survey for multiple subpopulations and for both continuous and categorical analysis variables (11 populations and 26 variables). Values of lambda were computed directly from the variance estimates. The lambda values were evaluated first by assessing the distribution of the lambdas and determining factors explaining the variation in the lambda values and then by assessing the effect of different lambda values on the point estimate and the variance estimates for the subpopulations and analysis variables. After values of lambda were identified, estimates were computed using the combined-sample weight and a second analysis assessed these estimates and the sampling variances.

These procedures resulted in a single value of lambda of 0.8606 being identified for the physician survey. This value achieved the desired level of sampling variances and simplified the processing of all estimates.

For the physician survey, the lambda value (0.8606) was estimated from the average of the medians for ten subpopulations of physicians. The evaluation of the effect of the lambda value (0.8606) indicated that the increase in the sampling variance will be between 1 and 3 percent for most subpopulations. For the larger populations, the sampling variances will increase by 4 to 5 percent. This increase in the sampling variance will be for populations that generally have smaller sampling variances. For income estimates, the average increase in the sampling variance

was approximately 10 percent, but in general, the sampling variances are sufficiently small that this increase is not likely to have substantive effect on analyses.

Section 6.3.7.1   Conceptual framework for combined-sample estimates

For computing survey estimates, Est(Y), combined across the two sample components, separate estimates can be computed for each sample component and combined using the equation

$$Est(Y) = \lambda \quad Y(Site) + (1 - \lambda) \ Y(Supp)$$

where Y(Site) is the survey estimate from the site sample, Y(Supp) is the survey estimate from the supplemental sample, and is an arbitrary constant between 0 and 1. For the sampling variance, V(Y), the estimate is computed using the equation

$$V(Y) = \lambda^2 \ V(Y(Site)) + (1 - \lambda)^2 \ V(Y(Supp))$$

where V(Y(Site)) is the sampling variance for the estimate from the site sample, and V(Y(Supp)) is the sampling variance for the estimate from the supplemental sample. Any value of lambda will result in an unbiased estimate of the survey estimate, but not necessarily an estimate with the minimum sampling variance. A lambda value producing a sampling variance at its minimum value results in the shortest confidence interval and, by implication, the most accurate point estimate.

A value of lambda can be computed in an optimal (minimum variance) sense as

$$\lambda = 1/V(Y(Site) / [1 / V(Y(Site)) + 1/ V(Y(Supp)]$$

$$= V(Y(Supp)) / [V(Y(Site)) + V(Y(Supp))].$$

In this case, the minimum variance is

$$V(Y) = [V(Y(Site)) \ V(Y(Supp))] / [V(Y(Site)) + V(Y(Supp))].$$

To compute the combined-sample estimate with minimum variance, survey estimates are derived by first computing the estimates for each sample component, computing a value of

lambda for each pair of estimates, and then combining the point and variance estimates. Although producing the minimum variance estimates, the process is computer intensive and results in some inconsistencies among estimates for percentages and proportions because of differing values of among levels of a categorical variable.

The alternative approach is to identify one or more values of lambda and compute combined-sample weights. To compute the combined weight for units (FIUs, persons or physicians) in the site sample,

$$WT(\text{Combined}) = \lambda \ WT(\text{trimmed site sample weight}).$$

For units in the supplemental sample,

$$WT(\text{Combined}) = (1 - \lambda) \ WT(\text{trimmed supplemental weight}).$$

After the combined-sample weight is computed, point and variance estimates can be computed directly using the SUDAAN survey data analysis software. The SUDAAN program code incorporates the estimation structure for the site sample and the supplemental sample as separate sets of strata.

Community Tracking Study Physician Survey

Sample Specification – Exclusions

<u>**Sample Specification – Exclusions**</u>

Physicans with the following primary specialities in the AMA Masterfile were excluded:

ALI	Allergy & Immunology/Diagnostic Laboratory
AM	Aerospace Medicine
AN	Anesthesiology
APM	Pain Management
ATP	Anatomic Pathology
BBK	Bloodbanking/Transfusion Medicine
CLP	Clinical Pathology
DDL	Clinical & Laboratory Dermatological Immunology
DMP	Dermatopathology
DR	Diagnostic Radiology
ETX	Medical Toxicology
FFP	Forensic Psychiatry
FOP	Forensic Pathology
HMP	Hematology
LM	Legal Medicine
MDM	Medical Management
MM	Medical Management
MPH	Public Health & General Preventive Medicine
NM	Nuclear Medicine
NP	Neuropathology
NR	Nuclear Radiology

OS	Other Specialty
PA	Clinical Pharmacology
PCH	Chemical Pathology
PCP	Cytopathology
PDR	Pediatric Radiology
PDT	Medical Toxicology
PIP	Immunopathology
PLM	Palliative Medicine
PMD	Pain Medicine
PP	Pediatric Pathology
PTH	Anatomic/Clinical Pathology
PTX	Medical Toxicology
R	Radiology
RIP	Radioisotophic Pathology
RNR	Neuroradiology
RO	Radiation Oncology
RP	Radiological Physics
SM	Sleep Medicine
SP	Selective Pathology
UM	Undersea Medicine
US	Unspecified
VIR	Vascular & Interventional Radiology

Physcians with the following primary specialties in the AOA membership file were excluded:

ALI	Allergy & Immunology/DLI
AM	Aerospace Medicine
AN	Anesthesiology
ANG	Angiography & Interventional Rad
AP	Anatomic Pathology
APL	Anatomic Pathology & Lab Med

APM	Pain Management

APPENDIX A

BBT	Bloodbanking/Tranfusion Medicine
CAN	Cardiothoracic Anesthesiology
CLP	Clinical Pathology
CP	Chemical Pathology
CY	Cytopathology
DPT	Dermatopathology
DR	Diagnostic Radiology
DUS	Diagnostic Ultrasound
EPI	Epidemiology
FOP	Forensic Pathology
FPS	Forensic Psychiatry
HEP	Hematology Pathology
IN	Internship
IPT	Immunopathology
IRA	Intraoperative Regional Anesthesiologist
LBM	Laboratory Medicine
LM	Legal Medicine
MMB	Medical Microbiology
NC	Nuclear Cardiology
NI	Nuclear Imaging and Therapy
NM	Nuclear Medicine
NPT	Neuropathology
NR	Nuclear Radiology
NRA	Neuroradiology
NV	In Vivo & In Vitro Nuclear Med
OBA	Obstetrical Anesthesiology
OE	Preventive-Occupational-Environmental Med
OS	Other Specialty
PA	Clinical Pharmacology
PAN	Pediatric Anesthesiology
PH	Public Health

PHP	Public Health & General Preventive Medicine

PMR	Pain Management-Rehab Medicine
PP	Pediatric Pathology
PRD	Pediatric Radiology
PTH	Anatomic/Clinical Pathology
PYM	Psychosomatic Medicine
R	Radiology
RET	Retired
RI	Body Imaging
RIP	Radiosotophic Pathology
RO	Radiation Oncology
RP	Radiological Physics
RT	Roentgenology
RTD	Diagnostic Roentgenology
SCL	Sclerotherapy
TR	Radiation Therapy
TX	Toxicology
TY	Transitional Year
UM	Undersea Medicine

APPENDIX B

Dear Colleague:

The Robert Wood Johnson Foundation is sponsoring a large-scale study on changes in the health care system and how these changes are affecting physicians, their practices, and the way they deliver medical care to their patients. Our goal is to provide results that will inform public and private leaders and enable them to make better policy decisions.

We would greatly appreciate your participation in this important endeavor. Under the direction of the Center for Studying Health System Change, interviews of physicians across the country are being conducted by the Gallup Organization, an internationally known survey research firm. The Center is a research organization funded by The Robert Wood Johnson Foundation to conduct this study and other studies of the health care system. All of the information you provide will be kept strictly confidential. It will be used in statistical analysis which precludes identification of individual respondents.

Your response is very important to The Robert Wood Johnson Foundation and to the success of the survey. We hope that we can count on your participation. We will be reporting the results widely. The following organizations support this study by urging their members to participate:

American Medical Association	American College of Physicians
Academy Osteopathic Association	American Psychiatric Association
American Academy of Family Physicians	American College of Surgeons
American Academy of Pediatrics	American Society of Internal Medicine

An interviewer from Gallup will be calling you soon to arrange an interview. We estimate that the interview will take about 20 to 25 minutes of your time and can be scheduled at your convenience. If you would prefer to contact Gallup directly at your convenience, you may call Michelle Stufing at 1-800-759-8789. For more information on the overall Foundation study, please contact Maureen Michael at 1-800-719-9419. Participating physicians will receive an honorarium of $25 as a token of our appreciation for their contribution of time to this important research effort. Thank you in advance for your cooperation.

Sincerely yours,

Steven A. Schroeder, M.D.

Sample letter sent to physicians selected for the first survey.

Route 1 and College Road East   Post Office Box 2316   Princeton, New Jersey 08543-2316   (609) 452-8701

Internet: http://www.rwjf.org
e-mail : mail@rwjf.org

APPENDIX B

Dear Colleague:

The Robert Wood Johnson Foundation is sponsoring a large-scale study on changes in the health care system and how these changes are affecting physicians, their practices, and the way they deliver medical care to their patients. Our goal is to provide results that will inform public and private leaders and enable them to make better policy decisions.

We would greatly appreciate your participation in this important endeavor. Under the direction of the Center for Studying Health System Change, interviews of physicians across the country are being conducted by the Gallup Organization, an internationally known survey research firm. The Center is a research organization funded by The Robert Wood Johnson Foundation to conduct this study and other studies of the health care system. All of the information you provide will be kept strictly confidential. It will be used in statistical analysis which precludes identification of individual respondents.

Your response is very important to The Robert Wood Johnson Foundation and to the success of the survey. We hope that we can count on your participation. We will be reporting the results widely. The following organizations support this study by urging their members to participate:

American Medical Association	American College of Physicians
Academy Osteopathic Association	American Psychiatric Association
American Academy of Family Physicians	American College of Surgeons
American Academy of Pediatrics	American Society of Internal Medicine

An interviewer from Gallup will be calling you soon to arrange an interview. We estimate that the interview will take about 20 to 25 minutes of your time and can be scheduled at your convenience. If you would prefer to contact Gallup directly at your convenience, you may call Michelle Stufing at 1-800-759-8789. For more information on the overall Foundation study, please contact Maureen Michael at 1-800-719-9419. Participating physicians will receive an honorarium of $25 as a token of our appreciation for their contribution of time to this important research effort. Thank you in advance for your cooperation.

Sincerely yours,

Steven A. Schroeder, M.D.

First letter sent to physicians

Route 1 and College Road East   Post Office Box 2316   Princeton, New Jersey 08543-2316   (609) 452-8701

Internet: http://www.rwjf.org
e-mail : mail@rwjf.org

APPENDIX B

Dear Colleague:

As a fellow physician concerned about current changes in American health care, I would like to ask you to take a few minutes to participate in a very important study that will guide some of the major health care policy decisions of our day. The study focuses on changes in the health care system and the practice of medicine and how these changes are affecting physicians and their patients. The Robert Wood Johnson Foundation (RWJF) is sponsoring this study conducted by the Center for Studying Health System Change, an independent research organization funded by the RWJF. The following major physician organizations have expressed their support for the study by encouraging their members to participate:

American Medical Association
Academy Osteopathic Association
American Academy of Family Physicians
American Academy of Pediatrics
American Academy of Pediatrics
American College of Physicians
American Psychiatric Association
American Society of Internal Medicine

One of the unique aspects of the study is our interest in how health care and the practice of medicine are changing at the community level. Your area is one of the communities randomly selected to be included in the study, and your participation is critical to its success. We will be able to accurately report on changes in your community, and across the nation as a whole, only if a representative sample of physicians participates. So far response to the survey has been excellent, but your participation is needed to ensure that all viewpoints are included. More information about the study is enclosed. If you have additional questions, please feel free to call Maureen Michael from the Foundation staff at 1-800-719-9419.

We are entering the final few weeks of the survey, so it is urgent to arrange a convenient time for you to complete a 15 to 20 minute telephone interview. The interviews are being conducted for this study by The Gallup Organization. Because we know you are very busy, we have arranged for an executive interviewer with experience interviewing physicians to contact you within the next few days to arrange an appointment to talk with you. We hope you will alert your office staff so they can put the call through to you or help arrange a convenient appointment time. If you would prefer to contact Gallup directly to set up an appointment at your convenience, you may call Michelle Stufing at 1-800-759-8789.

Your response is very important to the success of the study. Although we cannot compensate you fully for the time you spend participating in the study, you will receive $25 as a token of our appreciation for your assistance.

Thank you in advance for your cooperation.

Sincerely,

Steven A. Schroeder, M.D.

Second letter sent to physicians

Route 1 and College Road East   Post Office Box 2316   Princeton, New Jersey 08543-2316   (609) 452-8701

Internet: http://www.rwjf.org
e-mail : mail@rwjf.org