MINNESOTA ADULT TOBACCO SURVEY

Tobacco Use in Minnesota: 1999 to 2007

Minnesota Adult Tobacco Survey
MATS 2007 Methodology Report

September 2008

Minnesota Adult Tobacco Survey
MATS 2007 Methodology Report

ClearWay Minnesota[SM]

Blue Cross and Blue Shield of Minnesota

Minnesota Department of Health

Prepared by:

Westat
Rockville, MD

September 2008

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

Table of Contents (continued)

Table of Contents (continued)

List of Appendixes

Table of Contents (continued)

List of Appendixes (continued)

List of Tables

Table of Contents (continued)

List of Tables (continued)

Table of Contents (continued)

List of Tables (continued)

1. Introduction

The Minnesota Adult Tobacco Survey (MATS) is a comprehensive surveillance initiative designed to monitor progress toward meeting the goals of reducing tobacco use among Minnesotans. The major objective of MATS is to collect in-depth, public health surveillance data on the adult population of Minnesota, focusing on tobacco use and cigarettes in particular. MATS is the most comprehensive source of information about smoking prevalence, behaviors, attitudes and beliefs in the adult Minnesota population; further, MATS provides valid scientific data tracking the impact of comprehensive tobacco control efforts in Minnesota. MATS 2007 is the third survey in this ongoing surveillance initiative.

The MATS surveillance initiative and the three surveys—1999, 2003 and 2007—are directed by three partner organizations who lead comprehensive tobacco control efforts in the state of Minnesota: ClearWay Minnesota℠, Blue Cross and Blue Shield of Minnesota (Blue Cross) and the Minnesota Department of Health (MDH). Together, these three organizations formed the MATS 2007 Advisory Panel that selected Westat as the survey vendor for MATS 2007, made key decisions about survey design and provided oversight for the instrumentation, data collection, analysis and reporting of findings.

ClearWay Minnesota℠ is a nonprofit organization that strives to enhance life for all Minnesotans by reducing tobacco use and exposure to secondhand smoke through research, action and collaboration. ClearWay Minnesota serves Minnesota through its grant-making program, QUITPLAN® Services to help people quit smoking and statewide outreach activities. QUITPLAN Services helped more than 12,700 adult Minnesotans successfully quit tobacco use. ClearWay Minnesota designs and develops innovative statewide multimedia campaigns to inform the public of QUITPLAN Services and raise the awareness of the harm of secondhand smoke exposure. ClearWay Minnesota also works to build capacity and engage priority populations in reducing the harm that tobacco causes their communities. ClearWay Minnesota was created in 1998 when the state received $6.1 billion from its settlement with the tobacco industry and 3 percent, or $202 million, was dedicated

by the Ramsey County District Court to establish the independent nonprofit organization.

Blue Cross and Blue Shield of Minnesota is the largest health plan based in Minnesota, covering 2.9 million members in Minnesota and nationally through its health plans or plans administered by its affiliated companies. Prevention Minnesota is Blue Cross' unprecedented, long-term commitment to tackle preventable heart disease and cancers throughout Minnesota by addressing their root causes: tobacco use, exposure to secondhand smoke, physical inactivity and unhealthy eating. Prevention Minnesota is funded by Blue Cross' settlement proceeds from its landmark lawsuit with the tobacco industry, in which Blue Cross was a co-plaintiff with the state of Minnesota. Blue Cross and Blue Shield of Minnesota, a nonprofit corporation, is an independent licensee of the Blue Cross and Blue Shield Association. Blue Cross has provided stop-smoking programs for its members since 2000. Blue Cross also funds efforts to advocate for policy changes that help to reduce tobacco use and secondhand smoke exposure, works with high priority populations to raise awareness of the harm of tobacco use and promotes workplace health improvement.

The Minnesota Department of Health launched the first state-funded tobacco control program in the nation in 1985 with a portion of the proceeds from a cigarette tax. Since then, MDH has undertaken a number of tobacco control initiatives including participating as one of 17 American Stop-Smoking Intervention Study demonstration states, a national-level comprehensive tobacco control program sponsored by the National Cancer Institute. Funds from an endowment from the state's 1998 settlement with the tobacco industry were available to the department from 2000 through 2003 and were used to launch a comprehensive youth prevention initiative during that period. Currently, MDH works to reduce smoking through grants to reduce youth exposure to pro-tobacco influences, to create tobacco-free environments and to reduce tobacco related health disparities.

Through a competitive process, the Advisory Panel selected **Westat**, a leading health and social science research organization based in Rockville, MD, as the survey vendor for MATS 2007. Westat contributed technical expertise in sampling, weighting, and survey and analytical methods. With detailed direction from

ClearWay Minnesota, Blue Cross and MDH, Westat designed MATS 2007 and collected, analyzed and reported on MATS 2007 data. As a full-service vendor, Westat carefully reviewed the methodology of MATS 1999 and MATS 2003 and made recommendations to the Advisory Panel for adjustments based on the most up-to-date developments in survey research and study design.

The main components of MATS 2007 were as follows:

- **Sampling:** developing and drawing statistical survey samples that are representative of the Minnesota adult population and Blue Cross membership residing in Minnesota. The sample design called for a random-digit dialing (RDD) sample of the adult Minnesota population, combined with a stratified sample of Blue Cross members drawn from administrative records.

- **Questionnaire Development and Data Collection:** developing and administering a survey questionnaire that would obtain from the survey samples all the data items needed to support the larger health and tobacco-related missions of the sponsoring organizations. The questionnaire covered domains such as general physical and mental health, alcohol use, cigarette smoking and other tobacco use, smoking cessation, experience with healthcare provider smoking interventions, attitudes towards smoking, situational exposure to secondhand smoke in various settings, the effects of public and private policies and rules on smoking behaviors and perceptions, and demographic information. The questionnaire was administered using a computer-assisted telephone interviewing (CATI) system.

- **Survey Operations:** developing various operational procedures to support the administration of the questionnaire. These included telephone contacting rules and procedures that met or exceeded the standard requirements for the U.S. Centers for Disease Control and Prevention's (CDC) Behavioral Risk Factor Surveillance System surveys (BRFSS), along with supporting measures such as contact letters and an informational website.

- **Sample Weighting:** designing and creating sets of survey sample weights that can be used in analysis and reporting to make the sampled respondents' data statistically representative of the entire population they were designed to represent. Weights were based on the probability of selection into the

sample, as adjusted to selected available characteristics and counts of the adult Minnesota population. Separate sets of survey weights were developed for the RDD sample, the Blue Cross sample, and the combination of both samples that is to be used for analysis and reporting on the overall Minnesota population.

- **Tabulation and Analysis:** designing the various in-depth analyses of the survey data needed to support the sponsors' current and future programmatic, advocacy, public health, and tobacco-related health care delivery and insurance coverage activities, as well as developing the detailed analytical tools and specifications for tabulating and analyzing the data.

- **Reporting:** preparing an in-depth report profiling the adult Minnesota population in regard to the use, knowledge, attitudes, experiences, and plans surrounding tobacco/cigarette use, tobacco cessation, exposure to secondhand smoke, and other tobacco and health-related areas.

MATS 2007 data are to be used both to report the prevalence of tobacco use, exposure to secondhand smoke and related factors as of 2007 as well as to measure changes in these variables over time since MATS 1999 and MATS 2003. In this context, a critical objective for MATS 2007 was to maintain continuity with the previous two MATS surveys, primarily to support reliable tracking of population trends over time and to support inferential statements that observed significant changes over time reflect actual changes in the population and are not artifacts of differences in the survey design. Some changes in design are inevitable or unavoidable in large-scale surveys repeated over long time periods. Maintaining continuity in MATS 2007 was, therefore, a balancing act between adhering to past MATS instruments and methods and making desired or necessary improvements.

Comparability was also an objective for the design of the survey weights. Investigators from the University of Minnesota weighted MATS 1999 and MATS 2003 in accordance with generally accepted practices, such as CDC's BRFSS and other statewide tobacco surveys. These surveys generally post-stratify only on age and gender. In recent years, however, concern has grown among the research community regarding the representativeness of telephone survey samples, particularly in terms of educational attainment. Telephone surveys increasingly appear to be more likely to reach individuals with higher education attainment

(e.g., those with college degrees) than those with less education (e.g., those with a high school diploma or those who did not complete high school). While this phenomenon is not altogether new to survey research, the magnitude of the problem seems to have increased rapidly in the recent past. Because smoking and education status are inversely associated, the MATS Advisory Panel and Westat choose to include education as an adjustment factor for MATS 2007. To facilitate the most accurate comparisons between years of MATS administration, the data from MATS 1999 and MATS 2003 were reweighted to include educational attainment and race. Therefore, estimates from MATS 1999 and 2003 presented in MATS 2007 reports may vary slightly from estimates reported in previous publications.

1.2 Orientation to the Methods Report

This report constitutes the public documentation of general technical aspects of the MATS 2007 survey. It covers the sampling (Chapter 2), questionnaire development and data collection methodology (Chapter 3), the operational results of the data collection (Chapter 4), and the sample weighting (Chapter 5). Appendices include the MATS 2007 questionnaire, letters used in contacting the survey sample members, and the contents of a website that was created to provide information about the survey to potential respondents. The MATS 2007 analysis and reporting components are outside of the scope of this survey methods and appear as a separate, in-depth analytical report, titled *Creating a Healthier Minnesota: Progress in Reducing Tobacco Use* (Minneapolis, MN: ClearWay Minnesota℠, Blue Cross and Blue Shield of Minnesota, and Minnesota Department of Health; September 2008). This report can be found at www.mnadulttobaccosurvey.org.

The day-to-day development and conduct of MATS 2007 required many detailed, internal design, specification, and reporting documents and tools that are beyond the scope of this report. The Minnesota Adult Tobacco Survey 2007 Comparability Report[1] (Westat, July 2007) itemized the areas where MATS 2007 conformed to or diverged from MATS 2003, along with descriptions of the rationale for any

[1] Readers interested in more information about this report may contact Ann St. Claire, ClearWay Minnesota, at (952) 767-1416 or astclaire@clearwaymn.org.

differences and their potential impact on comparability as they might affect the findings of trends over time from MATS 2003 to MATS 2007.

1.3 Data Collection and Data Processing Timeline

Westat's work on the development of the MATS 2007 survey began in October 2006. Data collection concluded in June 2007 and final quality assurance checks of the data and post-coding of open-ended responses were completed by February 2008. Table 1-1 shows the timeline for the major activities of MATS 2007 from survey design through creation of the weighted data sets. Analysis and reporting activities are not included in this timeline.

Table 1-1. Timeline of MATS 2007 development, data collection and data preparation

Date	Task
Summer 1996 - 10/9/06	Select survey vendor, begin survey development (MATS Advisory Panel)
10/09/06	Hold kick-off meeting with Westat and Advisory Panel
10/09/06 - 1/29/07	Design, program and internally test MATS 2007 CATI questionnaire
10/9/06 - 1/29/07	Develop data collection protocols and supporting materials
10/9/06 - 2/6/07	Design samples, create sampling frames, draw and process sample for data collection (RDD & Blue Cross)
1/29/07 - 2/10/07	Conduct RDD pilot test and revise questionnaire
2/10/07 - 2/11/07	Telephone interviewer training
2/12/07 - 6/24/07	Telephone data collection
6/25/07 - 9/18/07	Final data quality assurance, post-coding, and RDD, Blue Cross, and composite weights for review & acceptance
10/07 – 2/08	MATS 2007 Advisory Panel review of weighting diagnostics; planning, design, and preparation of adjusted weights for MATS 1999 and MATS 2003.

2. Sampling

The MATS 2007 sample design utilized two sampling methods from two sample frames. One used a random-digit-dialing (RDD) statewide sample of adult Minnesotans, with a goal of conducting 7,500 interviews. The second used a sample of adult Blue Cross members living in Minnesota, with a goal of conducting 5,000 interviews. In an effort to improve the reliability of the statewide survey estimates, interview data from the Blue Cross sample were combined with the interview data from the RDD sample to create a single, weighted, combined dataset of 12,500 interviews, to be used for general reporting on the Minnesota population. In terms of the reliability of the Minnesota adult smoking prevalence rate to be estimated from the combined dataset, MATS 2007 was designed to detect a 2.5 percentage point difference between two estimates with 80 percent probability (power) at the 95 percent confidence level. The Advisory Panel initially established the Blue Cross sample size at 5,000 interviewed members. After accounting for the effect of the complex sample design on the statistical precision of the estimates, Westat determined that 7,500 completed interviews from the RDD sample needed to be added to the 5,000 from the Blue Cross sample for the combined dataset of 12,500 interviews to meet the reliability requirement,. The Blue Cross interview data were also separately weighted as their own dataset for Blue Cross to independently analyze and prepare reports about its membership. A single CATI questionnaire was used for both the RDD and Blue Cross samples.

2.1 Random-Digit Dialing (RDD) Sample

While some modifications were made, the basic RDD sample design for MATS 2007 followed the CDC BRFSS protocols, this mainly being that the sample of randomly generated telephone numbers was selected from banks of numbers that contain at least one "listed" telephone number. (A bank is a series of 100 consecutive telephone numbers with the same area code, exchange, and first two digits of the line number.) No sample was selected from banks with zero listed numbers.

Any differences from the BRFSS protocols that do exist revolved around issues of efficiency, that is, modifications were made in the sample design to improve the reliability of the estimates in a cost-neutral way. The methods employed are discussed in Section 2.1.1.

Beyond efficiency methods to complete 7,500 RDD interviews, the RDD sample selection process included methods to oversample young adults (aged 18-24) and African Americans. The oversampling of these two groups resulted from the desire to put additional focus on them in reporting the findings of MATS 2007. The specific techniques employed are discussed in Section 2.1.2.

A table with the RDD sample design projections and actual sample counts is provided in Table 2-1 in Section 2.1.3.

2.1.1 MATS 2007 Refinements of Basic RDD Sampling Methodology

2.1.1.1 Refusal Conversion Subsample

MATS 2007 used a disproportionate stratification technique that is based on refusals. For this procedure, Westat drew a larger sample of telephone numbers than would otherwise be selected to achieve the target number of completed interviews and then randomly pre-designated each telephone number to be assigned for household screener refusal conversion or not, if the screener was refused when the number was called. MATS 2007 used a subsampling rate of 60 percent to flag cases for which refusal conversion would be attempted, if the household member refused to answer the screener questions.

Westat fielded all the numbers subsampled for screener refusal conversion such that there was sufficient time for refusers to be worked during the field period. This technique is based on the observation that refusal cases comprise the majority of screener non-response in RDD surveys and substantial effort is expended to gain cooperation in households in which a member refuses to participate at the screener level. Drawing a larger initial sample and subsampling refusers (i.e., targeting the cases for refusal conversion) shifted some resources from the less productive, labor intensive task of refusal conversion to the more productive task of completing survey interviews. All subsampled screener refusal cases had sufficient time for the full refusal conversion protocol to play out (e.g., holding periods prior to initiating

refusal conversion, systematic recontact attempts after cases were re-released for conversion). In calculating the final weights, an adjustment was made so that the subsampled cases accounted for themselves as well as the cases that refused but were not subsampled for conversion (see Section 5.1.3). This subsampling was only employed at the household screener level; all refusals to the main interview received conversion attempts.

2.1.1.2 Within-Household Selection of Random Adult

The MATS 2007 sample design called for one adult to be selected at random from each household that was identified through the RDD screening process. To select an adult from within a household, a variation of the Rizzo method (for a detailed description of this selection method, see Rizzo et al. 2004), specifically a variation that allowed for the oversampling of young adults age 18-24 (see Section 2.2.1), was used to select an adult from within a sampled household. This variation involved obtaining a simple count of the number of adults age 18-24 and the number of adults age 25 or older. Once the number of adults in both of these age groups (18-24, 24+) was determined, the following logical hierarchy was applied:

1. If there was only one adult in the household, then that person was the sampled respondent.

2. If there were two adults and both fell into one of the two age groups, each had a 50% chance of being sampled, and the screener program utilized a computer-generated random number to determine whether the sampled respondent was the screener respondent or the other adult in the household. (This approach exploited the fact that the screener respondent must be an adult member of the household.)

3. If there were more than two adults and all fell into the same age group, the next-birthday method was used to select the sampled respondent.

4. If there were two or more adults and there was at least one in each of the two age groups, first the screening algorithm determined whether the 18-24 or 25+ age group was selected, using a computer-generated random number and a sampling rate that oversampled the 18-24 age group relative to the 25+ age group. Once the age group was selected, if there was only one person in that group, that person was the sampled respondent. If there were two or

more in that group, the next birthday method was used to select the respondent from within that group.

2.1.2 RDD Oversampling of Young Adults and African Americans

In addition to the general RDD sampling methods discussed in the previous section, MATS 2007 employed two sampling procedures to increase the reliability of two specific populations of interest: young adults and African Americans.

2.1.2.1 Oversampling Young Adults

Because of special interest in young adult Minnesotans age 18-24 years old, MATS sought to obtain a larger number of interviews with this age group than would result from their natural distribution in a random sample of 7,500 interviews. This desire was challenged by the fact that identifying and interviewing young adults through household screening processes has become increasingly difficult in RDD surveys, thought to be due largely to the dual factors of lower cooperation rates in this age group and reduced RDD sample coverage stemming from the increased number of cell-phone only households among young adults. Consequently, the MATS 2007 sample design oversampled young adults from the household screening process in households where both young adults (those 18-24 years old) and older adults (those 25 years of age and older) resided, referred to as "mixed-age" households. In such households, the young adult was selected 80 percent of the time. This particular oversampling rate of 80 percent was chosen to balance the need for extra young adult interviews with the increased variability in the weights resulting from both this oversampling and the oversampling of young adults in the Blue Cross sample (see Section 2.2.1). (Since the RDD and Blue Cross samples were combined for producing the statewide estimates, increased variability in the weights from both samples affected the reliability of the statewide estimates. The combining of these two files is described in Section 5.3.)

In contrast to MATS 2003, all older-adult-only households were retained in the sample in MATS 2007, rather than being subsampled as part of the protocol for oversampling young adults. The large majority of effort and costs would have already been expended at the point when the household screening was completed.

The relative inefficiencies of discarding some portion of older adult households were greater than the relative indirect statistical gains for the young adult sample.

The goal of this oversampling was to yield 1,300 RDD interviews with young adults, which, in combination with a projected yield of 900 interviews with young adults from the Blue Cross sample, would produce a total of 2,200 young adult interviews in the MATS 2007 combined data file. As described in Chapter 4, identifying and securing cooperation from the young adults in the RDD sample was more difficult than expected.

2.1.2.2 Oversampling African Americans

Because of a desire to conduct focused analyses of the African American population, MATS sought to obtain a larger number of interviews with this group than would result from their natural distribution in a random sample of 7,500 interviews. Hence as with the young adults, MATS 2007 sought to oversample African Americans.

The initial method for oversampling African Americans was to identify telephone exchanges with larger proportions of African Americans. The sample design placed all exchanges with greater than 15% Black/African Americans (based on Census 2000 population distributions by race) into a high-density stratum and oversampled phone numbers from these exchanges at three times the base sampling rate of the other exchanges.

The goal of this oversampling was to yield 450 RDD interviews with African Americans. As described in Chapter 4, identifying African Americans in the RDD sample was more difficult than expected, and the final yield of from the RDD sample was 290 African American interviews. Unlike young adults, who could be oversampled from the Blue Cross sample frame based on the existence of an age variable, race/ethnicity is not recorded in Blue Cross' records, so it was not possible to further oversample African Americans in the Blue Cross sample.

Close monitoring of the yield of completed interviews of African Americans during the initial weeks of the RDD data collection period resulted in a projection that the final total would be less than 200. This resulted in a decision to further supplement

the RDD sample with additional telephone numbers from exchanges with very high concentrations (greater than 25%) of African Americans. This effectively created three sampling strata defined according to the geographical concentration of African Americans: low concentration, high concentration and very high concentration. The additional supplemental sample resulted in an observed oversampling rate for the sampled phone numbers of 2.8 for the high concentration stratum and an oversampling rate of 8.3 for the very high concentration stratum, each compared to the low concentration stratum. The impact of this supplemental sample was to increase the final count of completed RDD interviews with self-identified Black/African American Minnesota adults to 290.

2.1.3 RDD Sample: Projected vs. Actual Counts

A random sample of telephone numbers was generated using the GENESYS Sampling System. These sample numbers were randomly formed into a number of "release" groups. Release groups allow for controlled, random release of sampled numbers, so that yields of completed interviews can be closely monitored and additional groups released to achieve the desired number of complete interviews, once the yield patterns become established. After an initial set of release groups was assigned to the MATS 2007 telephone interviewing operation, additional groups were released, as needed, to ensure the goal of 7,500 completed interviews was met. Also, a subset of the unreleased groups was used as the source for the supplemental sample of high density African American exchanges. See Table 2-1 for a breakdown of projected and actual RDD sample yields, which resulted in 7,532 completed RDD interviews. The statistics reported in the Actual column of this table reflect, in part, supplemental numbers added during data collection to address low yields in the African American RDD stratum. Full details on the RDD data collection operational results appear in Chapter 4.

Table 2-1. MATS 2007 RDD sample – design projections and actual

Sample outcomes	Design	Actual
Total sample drawn (including reserve)	98,100	116,010
Total sample needed	65,000	94,477
Electronically purged numbers (non-working & business)	32,950	56,170
Duplicates in super-high density African American stratum	-	10
Released to telephone interviewing	32,050	38,297
Non-residential/ineligible	7,830	13,167
Non-response	10,880	9,084
Unknown eligibility	3,470	6,021
Households screened	9,870	10,025
Extended interview non-response	2,370	2,493
Total extended interviews	7,500	7,532
Young adults	1,300	475
African Americans	450	290

See Chapter 4 for more detail on the outcomes of the RDD telephone interviewing operations.

2.2 Blue Cross Sample

The MATS 2007 Blue Cross population of interest consisted of members belonging to one of four Blue Cross and Blue Shield of Minnesota and Blue Plus plan types: Prepaid Medical Assistance Program (PMAP), MinnesotaCare, commercial plans and Medicare Supplemental plans. The MATS 2007 goal was to interview a total of 5,000 members, distributed proportionally across these plan types.

Beyond the goal of 5,000 total interviews, it was desired to increase the number of sampled young adult Blue Cross members. The reason for this is that, as noted above in Section 2.1.2, young adults are difficult to identify and reach through RDD methods. While the MATS 2007 RDD sample was designed to increase the yield of this group by oversampling them, obtaining a sufficient number of them to meet the MATS 2007 analytical objectives through RDD sampling alone would be costly, due to their low prevalence in the population compounded by their low response

rates. Since the comprehensive MATS 2007 sample design was based on combining the RDD and Blue Cross samples to produce statewide estimates, and since the Blue Cross members' ages are recorded in the administrative records, oversampling 18-24 year olds in the Blue Cross sample was a less costly way of supplementing the young adult interviews in the RDD sample. Because the different weights resulting from oversampling of the young adults within each of the two samples, and from the much larger weights of the RDD sample compared to the Blue Cross sample, it was necessary to strike a balance between obtaining as many young adult interviews as possible from the Blue Cross sample and avoiding swamping the combined file with too many interviews from the Blue Cross sample. The specific method for oversampling young adults in the Blue Cross sample and the sampling method in general are discussed below in Section 2.2.1.

A table with the Blue Cross sample design projections and actual sample counts is provided in Table 2-5 in Section 2.2.2.

2.2.1 Blue Cross Sample Frame and Sampling Stratum Definitions

2.2.1.1 Blue Cross Sample Frame Definition

The frame from which the Blue Cross sample was selected was a membership list of individuals covered by Blue Cross and Blue Shield of Minnesota or Blue Plus who lived in Minnesota . Such a sampling frame is often referred to as a "list frame". In addition to the four plan types listed above, the Blue Cross sample frame included members belonging to First Plan of Minnesota, a separately branded Blue Cross affiliate. First Plan members were implicitly included within the relevant main sample strata and flagged so they could be handled appropriately during data collection (e.g., sending them letters on First Plan letterhead, referring to their First Plan membership in the telephone introductory script). Due to contractual requirements, Blue Cross had to exclude members covered by products for state or federal employees and employees of one particular large private employer. Table 2-2 contains the MATS 2007 Blue Cross frame counts by plan type, before and after these exclusions. These counts reflect Blue Cross members living in Minnesota enrolled for coverage effective January 1, 2007.

Table 2-2. MATS 2007 Blue Cross frame counts by plan type

Plan type	Initial	Excluded	Blue Cross frame
Prepaid Medical Assistance Program (PMAP)	24,243	0	24,243
MinnesotaCare	33,615	0	33,615
Commercial plans	901,871	127,388	774,483
Medicare Supplemental plans	182,517	0	182,517
Total	1,142,246	127,388	1,014,858

2.2.1.2 Blue Cross Sampling Stratum Definitions

The Blue Cross sampling strata were classified by four plan types: (PMAP, MinnesotaCare, commercial plans, and Medicare Supplemental plans) and two age groups (18- 24 years and 25 years and older). Note that the Medicare Supplemental plan included a negligible number of individuals age 18 to 24 years old, so these cases were stratified with the 25 years or older stratum and are therefore counted as part of the 182,517 members in that stratum in Table 2-3. Table 2-3 contains the frame counts by stratum (after exclusion) for each of the seven MATS 2007 sampling strata.

Table 2-3. MATS 2007 Blue Cross frame counts by stratum

Plan type	Total	Age 18-24	Age 25+
PMAP	24,243	7,551	16,692
MinnesotaCare	33,615	7,286	26,329
Commercial plans	774,483	101,425	673,058
Medicare Supplemental plans	182,517	-	182,517
Total	1,014,858	116,271	898,587

2.2.2 Blue Cross Sample: Projected vs. Actual Counts

Discussion between Blue Cross and Westat led to the decision to sample Blue Cross members from the different plan types at the same rate within each age stratum, differing from MATS 2003, which sampled members at higher rates in some plan strata. The decision centered on the desire to increase the reliability of the

statewide estimates produced by the combined sample, while considering whether there was a need for Blue Cross to present estimates at the plan level, which had driven the need for a higher sampling rate from the smaller plan types in MATS 2003. Once it was determined that plan-level estimates were not a primary objective, it was feasible to sample each plan at the same rate within age stratum; this in turn improved the overall reliability of the resulting Blue Cross estimates, since the resulting sample weights would be closer in size for all the plans within the two age strata, especially as compared to MATS 2003. The sampling rate for each of the four age 25+ plan-type strata was 0.01129, while the sampling rate was 0.03387 for each of the three age 18-24 plan-type strata.

As discussed above, MATS 2007 oversampled Blue Cross members in the young adult strata. Based on discussions with ClearWay and Blue Cross, the final MATS 2007 sample design called for sampling young adult Blue Cross members at three times the base sampling rate of the older adults.

The Blue Cross sample was drawn using a stratified, systematic random sampling method. Within the each of the seven sampling strata, prior to drawing the sample, the frame was sorted in nested fashion by age (groupings defined by 5-year intervals within the age 25+ strata), gender, location (Twin Cities versus balance of Minnesota), and finally by a random number assigned to each record on the frame. Following the sort, a systematic sample was selected from each sampling stratum.

The drawn sample was large enough to allow for various forms of non-response while still achieving the target number of interviews. The sample size reflected conservative assumptions; for example, it allowed for a high proportion of the member records on the Blue Cross frame to have missing or out-of-date telephone numbers (as is frequently the case with administrative records), which would tend to lower the response rate, due to the inability to make contact with the sampled members. The MATS 2007 sample sizes by stratum appear in Table 2-4.

Table 2-4. MATS 2007 Blue Cross sample counts (total drawn) by stratum

Plan type	Total	18-24	25+
PMAP	576	332	244
MinnesotaCare	708	321	387
Commercial plans	14,346	4,466	9,880
Medicare Supplemental plans	2,678	0	2,678
Total	18,308	5,119	13,189

As in the RDD sample, the Blue Cross sample was randomly divided into smaller release groups, so that the interviews yielded from the sample released to data collection could be monitored and more groups released as needed to achieve the target of 5,000 Blue Cross interviews. The distribution of the seven strata in each release group was approximately proportional to their distribution in the total sample.

2.2.3 Blue Cross Sample and Operational Counts

Table 2-5 presents a breakdown of projected and actual Blue Cross sample yields, which resulted in 5,048 completed Blue Cross interviews. Full details on the Blue Cross data collection operational results appear in Chapter 4.

2.2.3.1 Blue Cross Telephone Number Look-up

MATS 2007 is a telephone survey. Since Blue Cross communicates with its members primarily by mail, their administrative records contain addresses for all members, but phone numbers only for some members; of the phone numbers on file, some may be out of date or otherwise unreliable. To increase the likelihood of reaching the Blue Cross sample by phone, Westat used two telephone number database vendors to confirm existing telephone numbers on the Blue Cross sample records, to obtain updated or additional telephone numbers when possible, or to obtain telephone numbers for those members who did not have a telephone number on their Blue Cross administrative record. Westat submitted the entire

sample to this look-up process, not only those sample records actually released to data collection.[2]

Table 2-5. MATS 2007 Blue Cross sample – design projections and actual

Sample outcome	Design	Actual
Total sample drawn (including reserve)	18,308	18,308
No telephone number available	915	828
Total sample needed	14,083	11,269
Non-callable sample – no telephone numbers available	705	467
Callable sample with telephone numbers[1]	13,378	10,802
Overlap telephone numbers with RDD[2]	130	166
Callable sample with telephone numbers, released to telephone interviewing[1,2]	13,248	10,636
Total non-response in telephone sample	8,248	5,588
Total completed interviews	5,000	5,048
Young adults	900	994

[1] See Section 2.2.3.1

[2] See Section 2.2.3.2

The first vendor found a telephone number match for 75.1 percent of the total drawn sample (including reserve), out of which 19.4 percent did not have a number in the Blue Cross records; this produced an additional 14.5 percent (.751 x .194) of the overall sample for whom contact could be attempted and who otherwise would have had to be classified as non-respondents from the outset. The second vendor matched a telephone number for an additional 6.2 percent of the sample that was not found by the first vendor, of which 17.1 percent did not have a number in the Blue Cross records. Using a second look-up vendor increased the selected sample for which data collection was possible by another 1.1 percent (.062 x .171). Combined, using the two vendors produced an additional 15.6 percent of the total

[2] Blue Cross member names, addresses, and phone numbers were provided to the telephone number database vendors without identifying them as Blue Cross members or MATS survey sample. The vendors keep such submissions confidential and do not use them for any other purpose other than to return the results of the look-up process to the requesting organization.

drawn sample (and the same percentage of the sample actually released) for which a data collection attempt would not have been feasible and which would have been immediately classified as non-response.

Counting sample with a phone number available from any source, 17,480 Blue Cross sample members out of the 18,308 were available to be assigned to data collection, of which 10,802 were actually needed. They appear as the "Callable sample with telephone numbers" in Table 2-5.

It is important to emphasize that, once sampled by a statistically valid process, any sample member for whom data collection cannot be completed for any reason (other than a determination of ineligibility according to the criteria defining the study population) must be considered as non-response, regardless of the reason, even if that reason is the lack of a phone number in a telephone survey.

If a sampled member had any associated contact number, either a number provided by the Blue Cross record or a number provided by the look-up vendors, it was assigned to telephone data collection. The number of sampled cases with one or more such phone numbers was 17,480, or 95.5 percent of the overall sample. There were 828 sample cases (or 4.5 percent of the overall sample) that did not have a contact number provided by any source. Among the sample actually released, there were 10,802 sampled cases with one or more phone numbers, or 95.9 percent of the sample actually released; there were 467 sample cases that did not have a contact number provided by any source, or 4.1 percent of the sample actually released. Since the latter were in the released sample but could not be assigned to data collection, they were immediately classified as non-response.

2.2.3.2 Blue Cross Sample Telephone Number Duplicates of RDD Sample Telephone Numbers

About a quarter of the adult Minnesota population are Blue Cross members. The Blue Cross and RDD frames therefore overlap, in that approximately 25 percent of the Minnesota population occur on both sample frames; as a result, a small proportion of the members drawn from the Blue Cross frame had telephone numbers that were also sampled as part of the RDD phone number sample. These are referred to as overlap telephone numbers. It is important to consider that the

Blue Cross sample is a sample of individual people, and a phone number is simply a way of reaching that person; in contrast, the RDD sample is initially a sample of phone numbers, and the number is a way to identify households and then select on person at random in the household. It would be theoretically possible and correct to call the Blue Cross member at his or her phone number to interview him or her as a sampled Blue Cross member, and to dial the same number as part of the RDD sample and select someone at random in the same household, who might be the same Blue Cross member or someone else in the household. The operational complexities and potential for confusion of such an approach would vastly outweigh the sampling purity that would result from this approach. To minimize respondent confusion and burden and, in some instances, sample duplication, MATS 2007 removed overlap telephone numbers from the Blue Cross sample. There were 287 overlap telephone numbers (less than 2 percent) in the overall Blue Cross sample, and 166 (1.5 percent) in the sample actually released. It is worth noting that, while these individuals were dropped from the Blue Cross sample, they still had a chance of being selected through the RDD sampling.

Westat removed these overlap telephone numbers from the Blue Cross sample rather than the RDD sample because they were a non-random subset of the RDD sample. Removing the overlap cases from the RDD sample would be akin to unilaterally and systematically removing any non-random subset of sample members with a certain characteristic; since insurance status correlates with smoking status, removing probable insureds from the RDD numbers would be even more biasing for MATS. On the other hand, the overlap telephone numbers are a random subset of the Blue Cross sample, since there is no known characteristic in the Blue Cross sample systematically associated with a member having one of these overlap telephone numbers, other than that they were also in the RDD sample. The RDD sample was itself a true random sample of the Minnesota population, as well as of any subset of that population, including Blue Cross members. Therefore, there is nearly no bias in dropping cases effectively at random from the Blue Cross sample, rather than dropping a non-random set from the RDD sample.

These cases appear as the "Overlap telephone numbers with RDD sample" in Table 2-5.

2.2.4 Comparability with MATS 2003 Telephone Number Look-up Fielding Protocol

According to information provided to Westat by ClearWay and Blue Cross, MATS 2003 made an assumption in regards to the Blue Cross sample members who had a phone number on their Blue Cross record, but no phone number returned from the single vendor look-up to which the MATS 2003 Blue Cross sample members were submitted. This assumption was that these "unlisted" Blue Cross members were similar to and therefore represented by those who did have a phone number returned by the look-up process. Since it is established that individuals without a listed phone number are typically harder to find or to secure a response from when found, the interview yields and response rates for the unlisted group would be expected to be lower than for those with listed numbers. In the interest of cost and operational efficiency, the MATS 2003 design dropped the unlisted group from data collection. For MATS 2007, Westat recommended including in the data collection sample every Blue Cross sampled member for whom it was feasible to attempt contact: those with a phone number only on the Blue Cross record, those with a phone number only from the look-up process, and those with a phone number from both sources (whether the same number or different numbers). This approach was likely to further minimize sampling bias and the Advisory Panel accepted this recommendation.

MATS 2003 treated both the cases with no phone number and the cases with a Blue Cross-provided phone number that lacked look-up confirmation as if they had not been sampled. As described below in Chapter 4, MATS 2007 adhered to the formal rule that a case cannot be considered as ineligible and dropped from the sample denominator solely because of barriers to making contact with the case, such as lack of a phone number or lack of a confirmed phone number. MATS 2007 included the latter cases in the sample assigned to data collection, while MATS 2003 did not. Neither MATS 2003 nor MATS 2007 could attempt data collection from the cases without any phone number, but MATS 2007 included them in the sample denominator for calculating response rates, while MATS 2003 did not include them in the denominator.

Table 2-6 compares the results on the sample of applying the MATS 2003 protocol relative to the more inclusive MATS 2007 protocol. It shows that 2,099, or 18.6 percent of MATS 2007 Blue Cross sample, would have been excluded under the MATS 2003 rules.

Table 2-6. MATS 2007 Blue Cross sample telephone number look-up result, broken out by rules for fielding telephone numbers under MATS 2003 and MATS 2007 protocols

Telephone number look-up result	Fieldable under MATS 2003 rules?	Fielded in MATS 2007?	MATS 2007 actual sample counts	Percentage of MATS 2007 sample
Blue Cross number and vendor number	Yes	Yes	7,383	65.5%
Blue Cross number, no vendor number	No	Yes	1,632	14.5%
No Blue Cross number, vendor number	Yes	Yes	1,787	15.9%
No Blue Cross number and no vendor number	No	No	467	4.1%
Total (% of total 11,269 MATS 2007 released sample)	9,170 (81.4%)	10,802 (95.9%)	11,269	100.0%

For methodological tracking and comparison purposes, Westat created a variable that flagged those cases fielded in MATS 2007 that would not have been fielded in MATS 2003 (those with the No-Yes pattern in Table 2-6). This variable allows for the subsetting of the MATS 2007 Blue Cross sample cases to emulate the MATS 2003 protocol in order to do a *pro forma* calculation of what the MATS 2007 Blue Cross survey response rate would have been under the MATS 2003 protocol. This pro forma response rate appears in Chapter 4; as expected, the exclusion of the "unlisted" cases (Blue Cross number, no look-up vendor number) produces a higher response rate than the actual MATS 2007 response rate that resulted from following the formal rules.

2.3 Sample Selection Quality Control

Westat and Blue Cross worked in concert when selecting the Blue Cross sample. Westat provided detailed specifications to Blue Cross describing how to prepare the sampling frame prior to actual selection of the sample. This included defining specific variables used to create stratification on the frame, sort ordering the frame by the strata and actual selection of records from the frame to create the sample. After delivery of the Blue Cross sample, Westat reviewed it to ensure expected counts by strata and characteristics within strata were satisfied.

In addition to the quality control of the sampling process, monitoring of yields for the RDD and Blue Cross samples, as well as the young adult and African American strata, occurred throughout data collection. Fine-tuning of sample releases and adjustments to the young adult and African American sampling approaches during data collection, as described below, were further measures designed to assure the quality of the sample.

3. Data Collection Methodology

3.1 Questionnaire Development

MATS 2007 required developing two questionnaires. The main questionnaire to be developed was the substantive survey instrument containing all of the questions for the MATS 2007 adult tobacco survey interview. In addition, the RDD sample needed a household screening questionnaire, some form of which is used in every RDD survey to identify households and then identify and sample people within the households. For brevity, the household screening questionnaire is generally referred to as the "screener" and the MATS 2007 adult tobacco survey questionnaire as the "extended" questionnaire; the latter term is used by many survey researchers because this stage of an RDD interview "extends" from the household screening interview. The same questionnaire was used for both the RDD and Blue Cross samples, and the term "extended" is used to refer to the adult tobacco survey questionnaire in general, even though the Blue Cross sample did not involve a screener.

3.1.1 MATS 2007 Questionnaire

The MATS 2007 Advisory Panel began the process of designing the MATS 2007 instrument in the summer of 2006, by reviewing the MATS 2003 instrument and proposing items to be added, eliminated or reworded. Applying an iterative, consensus approach, they worked through various versions, adding items to address new research questions or provide further information about previous research questions. The proposed changes reflected the current research agenda of MATS 2007, experience with the utility of MATS 2003 data, and the need to eliminate some items to accommodate new items. The final decision to eliminate a question or panel of questions usually reflected a general consensus that the eliminated items were of interest in the past but not in the present, or were of lesser importance, given the need to obtain different information in MATS 2007. Items that appeared in the MATS 1999 or MATS 2003 questionnaire but were never used in the resulting analyses were likewise eliminated.

Westat began working with the Advisory Panel to refine and finalize the design of the questionnaire in October 2006. While some items remained to be added or eliminated in mutual discussions, Westat focused on working with Advisory Panel representatives on question wording, response category selection and wording, question flow and ordering, and optimizing the design for telephone interviewing.

As noted in Section 1.1, maintaining continuity with the previous two MATS surveys was a critical objective. However, changes in questionnaire design are desirable or unavoidable in large-scale surveys repeated over long time periods, due to the emergence of new issues or phenomena in regard to tobacco control, epidemiology, treatment, and education; scientific advances; altered focus on the part of the researchers, administrators, and practitioners who use the time series data; and the impacts of real world occurrences, such as political forces, actions of the tobacco industry, funding limitations, or social factors. Maintaining continuity in the MATS 2007 questionnaire was a balancing act between absolute conformity and making desired or necessary improvements.

Examples of substantial changes made in the MATS 2007 instrument include the elimination of questions that delved into certain details of how stop-smoking medications or nicotine replacement therapy were used, the helpfulness of quit aids used by successful quitters, and detailed information about insurance coverage. The elimination of these questions was consistent with the changing relative importance of MATS 2007 research questions to ClearWay and Blue Cross. For example, although MATS 2007 included research questions about quitting smoking, the addition of research questions about the relationship of tobacco control media messages, smoke-free policies, price increases, and smoking bans to quitting behavior meant that detailed items about the types, use, and helpfulness of quit aids, while of interest, were less critical than in MATS 2003 and were therefore among those eliminated to accommodate the desired new items.

The questions that were added to the MATS 2007 instrument allowed exploration of the new research questions by asking about media messages, the effect of smoke-free policies at work and home, and buying cigarettes. In addition, some new questions were refinements of previous constructs that allow ClearWay and Blue Cross to more closely examine issues related to use of alternative tobacco products

(hookahs, chewing tobacco), the context of smoking, and awareness of free sources of help in quitting. Like the questions that were eliminated, the added questions were consistent with the changing relative importance of previous research questions to ClearWay and Blue Cross and with the new MATS 2007 research questions.

Occasionally, retained questions and response categories were reworded to make their meaning clearer to respondents. Each proposed change was carefully considered by the design team, who weighed the improvements to meaning and clarity against the effect on data continuity. Based on experience in MATS 2003 or the expert judgment of the design team, some response categories for some questions were eliminated or collapsed, or new categories added. Another type of change was a combination of adding and rewording questions and altering response categories. In a few instances, breaking one question into two questions or inserting a set-up question before or a follow-on question after an existing question improved the clarity of meaning. For example, the MATS 2003 instrument asked the following two questions about work:

> [Q43A] While working at your job, are you indoors most of the time?
>
> [Q43B] What best describes where you work for money? Is it a classroom, hospital, office, your home, a plant or factory, store or warehouse, restaurant that does not serve alcohol, restaurant that serves alcohol, bar, vehicle, or some other environment?

In MATS 2007, the second question was broken into two questions in which the response categories were divided respectively into indoor or outdoor locations, to simplify the respondent's ability to form a response to the question. A skip pattern added to the first question determined which set of responses (indoor or outdoor) was offered to the respondent. While such changes could improve data validity, there was always the consideration that such improvements could produce observed changes from MATS 2003 as an artifact of survey design, in this case, the restructuring of the questions, even though the logical content is itself was unchanged. In each such instance, the design team considered the possible size of

the positive and negative effects before deciding whether to implement the change. In a few instances, the ordering of questions was changed to clarify the meaning or clarifying language was inserted prior to the question. When the changes were adopted, it was because they were deemed necessary and with sufficiently low impact on the comparability of responses of the affected questions (e.g., introducing new topical sections before retained questions) or likely to improve data validity at minimal risk to continuity.

The final MATS 2007 questionnaire appears as Appendix A of this methods report. In addition to developing this interview script format of the questionnaire during the design process, Westat also developed two tabular formats for documentation and quality control. One was a detailed table showing the skip patterns for every question for every smoking status or other criteria that affected skip patterns, which appears as Appendix B. The second was a detailed table crosswalking and documenting every question or response category added, deleted, or changed from MATS 2003 to MATS 2007, along with an assessment of its possible impact on data comparability between MATS 2003 and MATS 2007. This table is incorporated in the Minnesota Adult Tobacco Survey 2007 Comparability Report which provided further details of the changes and additions that resulted in the MATS 2007 questionnaire.

The final MATS 2007 questionnaire covered domains such as general physical and mental health, alcohol use, cigarette smoking and other tobacco use, smoking cessation, experience with healthcare provider smoking interventions, attitudes towards smoking, situational exposure to secondhand smoke in various settings, the effects of public and private policies and rules on smoking behaviors and perceptions, and demographic information.

Westat developed detailed specifications to program the MATS 2007 questionnaire as a CATI survey instrument. The programming specifications are embedded in the MATS 2007 instrument included as Appendix A.

3.1.2 MATS 2007 RDD Household Screener

Appendix C is the MATS 2007 RDD household screener instrument. Westat developed this approach to the screener to implement the Rizzo method for RDD respondent selection and also to support the oversampling of young adults.

The final version of the screener in Appendix C represents a modification of the version used at the start of data collection. The questions about the presence of young adults in the household and the wording to communicate to the screener respondent the selection of a young adult were slightly modified during data collection when the yield of young adult respondents – always low in RDD studies – was even lower than expected. Investigation revealed that, during their actual interview, some respondents selected as young adults were reporting an age older than 25 years old, indicating either that the screener respondents had misunderstood the questions about the ages of household members or misunderstood the request to speak with the specific selected respondent. During data collection, the screener was revised by clarifying the wording for certain questions and instructions and, most importantly, also by adding confirmation questions about the reported age composition of the household.

3.1.3 CATI Questionnaire Programming and Testing

Programming of the CATI questionnaire was carried out by Westat's CATI programming team, led by a senior CATI systems analyst. Testing of the programmed instrument was performed by the programmers, by an independent testing department at Westat, and by questionnaire designers from Westat and the MATS 2007 Advisory Panel. The several levels of testing revealed a few items that required correction and a few items that resulted in minor changes to the instrument design and specification.

3.2 Pilot Test

3.2.1 Background

Between January 29 and February 10, 2007, Westat conducted a pilot test of the RDD survey, including the MATS 2007 questionnaire, the RDD household screener, the within-household sampling procedures, the interviewer scripts and telephone

contact procedures, and the handling of the cases in the CATI system's automated scheduling and case management system. By design, the pilot test did not include the Blue Cross member sample; since the Blue Cross sample used the same MATS 2007 questionnaire as the RDD sample, the RDD pilot test would address most of the aspects of interviewing the Blue Cross sample, except for the different initial contact procedures for the Blue Cross member sample. The pilot test objectives were live field testing of the:

1. Programming of the CATI questionnaires;
2. MATS 2007 questionnaire's suitability for administration by interviewers;
3. Respondents' comprehension of the questions and their ability to provide answers;
4. RDD Screening Questionnaire, screening rules and procedures, and respondent selection; and
5. Assumptions for RDD number sample yields, screener and individual cooperation rates.

3.2.2 Pilot Test Operations

The pilot test had a goal of 100 completed interviews. An initial sample of 1,100 RDD numbers was assigned to data collection. As discussed below, one of the pilot test findings was that residential number yields for the Minnesota RDD sample were lower than typically experienced in national samples, so an additional 1,000 numbers from reserve sample were added to the test sample during the test period. This also alerted the statistical and data collection managers to anticipate releasing more numbers from the reserve sample in the full survey than originally projected.

The pilot data collection employed substantially all of the data collection procedures to be implemented in the full survey, with two planned exceptions. Because the data collection period of the pilot test was brief and the primary objective was to test the instrument, the pilot test purposely did not carry out the refusal conversion protocol for either the household screener or the extended interview, as planned for the full survey. The pilot test sample also did not receive any of the supporting letters (advance letter, non-contact letter, and refusal

conversion letter) that the actual survey sample received (see Section 3.4.1 for a full description of these letters).

3.2.3 Pilot test Interview Monitoring and Interviewer Debriefing

During the pilot test, Westat's telephone supervisory staff conducted live monitoring of the interviews. Monitors could hear both sides of the conversations and see on their computer screens a live, mirrored version of the interviewer's actual CATI screen. During regular survey operations, the monitoring is conducted as a quality assurance measure of the interviewer's following of the data collection protocols, correct reading of the questionnaire text, handling of questions and problems, and entry of responses. Since highly experienced interviewers were assigned to the pilot test, the monitors were able to focus on the aspects of the data collection design that were being subject to testing: whether the procedures worked as planned, whether the questionnaire wording and flow supported clear administration by the interviewers, and whether the respondents had any general difficulties in understanding the questions or formulating an answer. The monitoring produced no reports of general problems along these lines. Interviewers were able to handle questions that some respondents asked about the survey or specific questions, based on their training and utilizing the set of Frequently Asked Questions developed for them by the survey managers.

When about three-fourths of the interviews were completed, Westat's telephone operations manager and the MATS 2007 project director held a focused debriefing session with two interviewers and one supervisor. It consisted of two parts: an opening discussion in which project managers asked the interviewing staff to provide their unprompted reactions to the questionnaire design and the interviewing experience, followed by a structured set of questions for them about specific questions and issues.

The statisticians examined the various yields of the pilot test, including prevalence of young adults in the screened households and the actual sampling rates for young and older adults. All of these factors matched those specified in the statistical sampling design, allowing for the sampling error of the small pilot test sample.

The interview length timings also came within the desired upper limits designed for different types of respondents (smokers, former smokers, never smokers).

The MATS 2007 pilot test revealed that the design of the MATS 2007 questionnaire successfully achieved its objectives, in terms of obtaining the desired information, being clear and unburdensome to respondents, and readily administered by interviewers. The live test also confirmed that the CATI instrument performed correctly as to flow and data capture, as intended by the design and as previously verified by beta testing. Perhaps most importantly, the test showed that the overall design of MATS 2007 – interviewer training, calling procedures, the RDD screening, explaining the purpose of MATS 2007, identifying and obtaining cooperation from selected individuals, and successfully taking them through the MATS 2007 questionnaire – was well considered and feasible in a real world setting.

3.2.4 Pilot Test Findings

As a result of the pilot test, ClearWay, Blue Cross, MDH, and Westat agreed to a few minor changes to improve specific areas. As described more fully in the Minnesota Adult Tobacco Survey 2007 Pilot Test Report[3] (Westat, 2007), the principal questionnaire changes were:

- Dropped mention of ClearWay as a sponsor in the RDD introduction
- Changed the cycling order of questions G1 – G4, about use of health care providers and receipt of various forms of quit-smoking assistance from them
- Shortened wording for H2, about frequency of experiencing advertisements about quitting smoking and the dangers of secondhand smoke
- Allowed interviewers to accept responses without reading the entire list in questions where the respondent can provide an explicit answer without hearing the entire list, such as H14, which asks about type of work location.

Westat presented the specific findings and additional details about the pilot test in the MATS 2007 Pilot Test Report.

[3] Readers interested in more information about this report may contact Ann St. Claire, ClearWay Minnesota, at (952) 767-1416 or astclaire@clearwaymn.org.

3.3 Interviewer Recruitment and Training

Westat assigned interviewers from its current staff of interviewers located within its telephone call centers and recruited additional staff as needed to supplement current staff. The additional staff were located either in call centers or worked from their homes. At the request of ClearWay and Blue Cross, Westat focused on Minnesota and nearby states in recruiting new at-home interviewing staff whose first assignment would be MATS 2007.

All interviewers received two waves of training: general interviewer training and MATS-specific training. (Current interviewers had previously received the general interviewer training; newly recruited interviewers received both). The training was self-paced and self-administered, with the interviewers working though Web-based self-study modules. Before an interviewer could progress to the next module, he/she had to pass an assessment on the module just completed, with a 100 percent correct score. Interviewers could review content until they were able to attain 100 percent correct on all assessments.

The general interviewer training (GIT) modules covered topics such as:

- The concept of data and social science research, and the role of the interviewer in this research process
- Principles and tenets of standardized interviewing, the elements of professionalism, and the use of the CATI system
- Concept of a scientific sample and the importance of probability sampling; the various ways data can be collected: in-person, telephone, Web, observation, medical measurement, etc.
- Survey design, administration, and respondent contact procedures.
- Standard set of call disposition codes
- Importance of interviewer neutrality, verbatim question delivery, and exact recording of responses as central to standardized interviewing

- Projecting professionalism and expert knowledge of the survey as key characteristics in securing respondent cooperation
- Listening skills and speaking skills
- Voluntary nature of survey participation, informed consent, and confidentiality

Training in use of the CATI system employed an interactive, self-administered, computer-based tutorial. Each interviewer moved through a series of topics, such as instruction on logging onto the CATI system, using the keyboard, the mouse, and special CATI commands. At this point interviewers learned the keys and commands for entering data and handling situations outside of the automated flow of the CATI questionnaire. Also included in this session was practice in the coding of contact results. Trainees experienced recorded replications of common contact situations and learned the proper coding techniques through presentation and practice. A follow-up test was administered to evaluate mastery of the contact procedures. The interviewers who were considered for the MATS assignment and to receive the MATS-specific training were limited to those who achieved a perfect score (100%) on this test.

Westat based the MATS-specific training on the BRFSS training protocol as described and demonstrated at the BRFSS Web site http://www.cdc.gov/brfss/training/interviewer/index.htm. This protocol emphasizes the presentation of questions, instructions to interviewers on administering the questions, and the rationale for the questions, so that the interviewers understand the importance of following the protocol.

The MATS-specific training was conducted in the same way as the general interviewer training, including the requirement to score 100 percent correct on each module's evaluation assessment before proceeding to the next training module. The modules focused on:

- MATS questionnaire items and the flow of the MATS questionnaire, including terms and definitions
- RDD screening process

- Contact scripts, locating and verifying the Blue Cross sample; obtaining new phone numbers to pursue Blue Cross sample members not found at the current available phone number
- Handling of problem situations and the use of the prepared, standard responses to frequently asked questions.

Training instructors and team leaders were available to assist interviewers should they encounter difficulty with a particular training concept or module. MATS trainers communicated with interviewers through instant messaging, training-based electronic bulletin boards, email, and telephone calls. After interviewers completed all of their training modules, they were teamed with a partner to conduct practice interviews with each other (role play). During the role play sessions each interviewer was monitored and coached to assess and enhance their interviewing skills. Once the supervisor determined the interviewer demonstrated the appropriate command of the interview and study materials, the interviewer was permitted to conduct actual MATS interviews.

3.4 Communications with Sample Members

There were a variety of methods used to communicate with the MATS 2007 sample prior to and during data collection. These included a variety of letters, an informational website, and several contact numbers that potential respondents could call for information or other purposes. These tools were designed to improve response rates and provide information to sample members or the general public about the survey.

3.4.1 Letters

3.4.1.1 RDD Letters

MATS 2007 developed three different types of letters sent to members of the RDD sample. These were an advance letter, a non-contact letter and a refusal conversion letter. All RDD letters were printed and mailed using letterhead of the Minnesota Department of Health and signed by the Director of the Center for Health Statistics.

- The **advance letter** was mailed to each address that could be associated through a listed number database with a sampled RDD number that was in

the released sample. It was addressed generically to "The Household at…" the matched address. Its purpose was to inform the household of their possible inclusion in the survey, give them information about it, stress the voluntary and confidential nature of the interview, urge their participation if selected, provide them a reference to the MATS 2007 web site (described in Section 3.4.2) and the phone number at MDH that they could call for more information. The letters cited all of the survey sponsors (ClearWay, Blue Cross, and MDH).

- The **non-contact letter** was a variant of the advance letter. It was mailed when, after repeated calls to an RDD number, no contact had been made that would allow determination of whether or not the phone number belonged to a residence. Like the advance letter, the non-contact letter could be sent only for phone numbers that had been matched to an address. The non-contact letter was designed to get through to people who may have been screening calls through Caller ID, who may have had a phone line to which no phone was connected, or similar reason why contact could not be made. The letter stressed the importance and legitimacy of the survey and urged the recipient to respond to calls from Westat or to call Westat's toll-free number.

- The **refusal conversion** letter was mailed when a phone number had been established as belonging to a household but the members refused to participate in the household screening interview. Like the advance letter, the non-contact letter could be sent only for phone numbers that had been matched to an address. The refusal conversion letter was designed to persuade the household to participate in the screening and then in the extended interview. It contained much of the information included in the advance letter, with additional emphasis on the importance of the recipient's participation.

3.4.1.2 Blue Cross Letters

A similar set of letters with similar contents and purposes was sent to the Blue Cross sample members. Blue Cross sample letters were printed and mailed using the letterhead of Blue Cross and Blue Shield of Minnesota and Blue Plus and were signed by the Vice President and Medical Director of Population Health. Because a small number of individuals in the Blue Cross sample belonged to a separately branded Blue Cross affiliate called First Plan of Minnesota, a second version of the letters for the Blue Cross sample was created. The two versions were nearly

identical, except that the First Plan members received letters on First Plan letterhead, the sponsorship included a reference to First Plan, and the letters were signed by the First Plan Medical Director

While essentially the same as the RDD letters, the Blue Cross letters had a few differences of note.

- Since the Blue Cross sample consisted of specific, individual Blue Cross members, and Blue Cross has current address information for all its members. Each type of letter was sent to the sampled member by name, at the address of record.
- Since there was no household screener, the refusal conversion letter was mailed when the sampled respondent (or another intermediary in his or her household) refused participation in the survey.
- Because the sample consisted of plan members, the Blue Cross letters contained additional assurances that the decision of whether or not to participate in the survey would have no effect on their health insurance coverage, premium costs, or relationship with their provider; that Westat would not furnish their identity back to Blue Cross; and that the information provided in the interview could not be linked to their Blue Cross records. All three letters provided the recipients with the name and phone number of a Blue Cross representative they could call if they wished to verify the survey or to decline participation in the survey upon receipt of the letter. (As further discussed in section 3.4.3, sample members who called Blue Cross to decline participation were immediately removed from the active phone sample and classified as final refusals.)

Copies of the nine letters appear in Appendix D.

3.4.2 MATS 2007 Informational Website

The MATS 2007 Advisory Panel and Westat web designers developed an informational web site to provide sample members and potential respondents with a set of brief, simple, and clear informational points about the survey. Its purpose was to encourage participation among selected respondents, enhance the perceived legitimacy of the survey, and answer questions potential participants might have. Legitimacy was enhanced by the visibility of a website accessed through a link on

the Minnesota Department of Health's official website, at www.health.state.mn.us/2007HealthStudy. This link took users to a website created and maintained on Westat's web server, with the address www.MNHealthStudy.org. The latter was a domain name temporarily acquired for use with MATS 2007.

The main webpage provided a brief overview of the survey, with menu links to four subpages covering the topics:

- How participants are selected.
- How the survey works.
- Frequently asked questions.
- Sponsoring agencies and contact information.

In the letters, web pages, and web domain names, MATS was characterized as a health study with an emphasis on tobacco rather than exclusively as a tobacco survey. This was designed to avert non-smokers from a disinclination to participate due to a perceived lack of relevance to them and to mitigate smokers' possible perceptions of persistent focus on them by media, government, and the health care community.

The contents of the MATS 2007 informational website appear in Appendix E.

3.4.3 Toll-free Numbers and Contacts Provided by Westat, Blue Cross, and Minnesota Department of Health

Westat operated a toll-free number that MATS 2007 sample members could call to obtain information about the survey. Westat's inbound call center answered the toll-free line and either responded with the requested information or referred the caller to an assigned contact person at Blue Cross or MDH. Blue Cross provided a toll-free number and the name and extension a person to contact; MDH provided the name and direct line of a contact person. These various numbers and contacts were printed in the appropriate letters, were available on the website, and were provided upon request by the telephone interviewers.

In addition to supporting general informational requests, the Blue Cross toll-free line served as the channel for Blue Cross sample members who chose to inform Blue Cross in advance of being called by the interviewers that they did not wish to participate in the survey. The advance and other letters informed them of their right to "opt out" of the survey. Efforts to interview Blue Cross members who opted out by this means ceased immediately when they notified Blue Cross.

3.5 Data Collection Confidentiality Procedures and Protection of Human Subjects

All Westat staff are bound by strict confidentiality and privacy rules and procedures, designed to prevent deliberate or inadvertent disclosure of the identity or survey data of anyone belonging to a data collection sample. All Westat staff are trained in the relevant protocols, covering oral, electronic, or printed disclosure, and in the techniques to safeguard such information in all these forms. As a condition of employment, they are required to sign a pledge of confidentiality laying out these requirements.

Both center-based and home-based interviewers were required to adhere to the same set of confidentiality rules and procedures, including the signing of a strict pledge of confidentiality. Home-based interviewers were subject to further requirements, in terms of working from a segregated office space within their home environment and outside of the presence of anyone else in the household. Using the web-based interface, all data collected by the home-based interviewers were entered in real time into the central survey database maintained within Westat's physical facility and behind Westat's software firewall. Sample identifying information, questionnaire text, and response data were only visible on the interviewers' screens; no data could be copied or saved electronically or printed locally.

All sample and survey data were maintained on Westat's secure, password protected network, with access to MATS-related data limited to approved MATS 2007 project staff.

The MATS 2007 survey questionnaire, data collection, and data security plan were reviewed and approved by the Westat Institutional Review Board (IRB), a specially constituted review body established to protect the welfare of human subjects recruited to participate in biomedical and behavioral research. Westat's responsibilities are detailed in the regulations concerning human subjects protections and the Multiple Project Assurance (MPA) granted by the U.S. Department of Health and Human Services, Office for Protection from Research Risks, Division of Human Subject Protection.

Westat's general confidentiality procedures are designed to comply with applicable requirements of state and federal law relating to Protected Health Information (PHI), including the Health Insurance Portability and Accountability Act of 1996 (HIPAA). PHI and HIPAA apply to health information contained in health records; collected survey data is covered by other federal statutes and is subject to the oversight of the Office of Protection from Research Risks. Westat's contract with Blue Cross required that Westat appropriately safeguard any PHI made available to Westat. Westat received only name, contact information, and three sampling quality control variables (age, gender, plan type). At no time did Westat have access to Blue Cross' records; Blue Cross drew the sample in accordance with Westat's specifications. At no time did Westat have access to any information in regard to medical history, health status, or insurance claims.

3.6 Data Collection Operations

Data collection for the RDD sample occurred between February 12 and June 23, 2007, and for the Blue Cross sample between March 16 and June 24, 2007. Calling took place from 9 AM to 9 PM weekdays, 10 AM to 6 PM Saturdays, and 2 PM to 9 PM Sundays (all times Central time). Consistent with standard operational practice for personal telephone interviews of individuals in their homes, the majority of calls occurred between 6 PM and 9 PM weekdays and throughout the weekends, in order to optimize the amount of effort applied when people are most likely to be found at home.

3.6.1 Calling Procedures

3.6.1.1 General Case Handling and Contact Procedures

MATS 2007 telephone procedures applied a hierarchical approach to case management. This includes making cases available for call attempts based on the current status of the call. Those cases for which an exact appointment was scheduled had the highest priority, followed by those with a general call back time (e.g., information that "evenings are best" to reach the desired person), those that had been called previously without human contact and lastly cases that had never been called. Consistent with BRFSS protocol, at least 15 call attempts were made to each sampled number, unless the number was determined to be non-working or assigned to a business. Until contact was established, the CATI scheduling system automatically spread out the calls across various times of day and various days, including some weekdays and some weekend days.

As described in section 3.6.3, a second effort was made to convert refusals to the RDD household screener for those households that had been randomly flagged in advance for refusal conversion, as well as all refusals of the extended interview, except for those few characterized as "hostile" (or "adamant") refusals.

3.6.1.2 Supplemental Calling Procedures

Once data collection was in progress, MATS 2007 implemented several measures designed to improve response rates and increase the number of completed interviews obtained from the released RDD and Blue Cross samples. The most significant of these was the decision to re-activate cases that had been coded as final non-response because they had reached the maximum number of call attempts according the BRFSS/MATS protocol. Periodically during the second half of the data collection period, Westat re-activated such cases for another round of calls.

Late in the data collection period, one weekday and one weekend day were designated to allow every call made on those days to ring more times than the normal limit allowed before the call was coded as a no-answer. This was designed to reach people who may have needed the extra time to reach the phone or who purposely allow their phone to ring for a certain number of rings before answering it, e.g., to avoid telemarketers.

In order not to badger households, the MATS calling protocol allowed for only one answering machine message to be left at a phone number prior to the point when live contact was made. Towards the end of data collection, a second answering machine message was left at any phone number where live contact still had not been made.

3.6.1.3 Locating and Verifying Blue Cross Sample Members

Since MATS 2007 was a telephone survey, any sampled Blue Cross member who could not be located and reached by telephone was considered a non-respondent. Section 2.2.2.1 described the limitations of the telephone number information contained in the Blue Cross records and the submission of the Blue Cross sample to two vendors of listed telephone number information. Westat submitted to these vendors the names and addresses Blue Cross provided at the time of sampling.[4] The look-up process from each vendor could produce the same phone number as on record, a different phone number than on record, a phone number where none was on record, or no number where no number was on record.

To increase the likelihood of reaching the Blue Cross sample by phone, Westat assembled all unique phone numbers obtained from any source into a prioritized list before beginning calling, placing a number found by the look-up process at a higher priority than the Blue Cross number, on the assumption that it represented more current information. Thus, there were one or two numbers available to be tried for cases assigned to telephone data collection; those without any number from any source were withheld from data collection and classified as non-response.

The prioritized list was placed into the sample case management records in the CATI case management database and the CATI call scheduler automatically dialed the number with the first priority during each attempt to contact the sample member, until one of three possible scenarios was encountered.

[4] Blue Cross member names, addresses, and phone numbers were provided to the telephone number database vendors without identifying them as Blue Cross members or MATS survey sample. The vendors keep such submissions confidential and do not use them for any other purpose other than to return the results of the look-up process to the requesting organization.

1. If contact with someone at that number revealed that the sampled member did not live at the location reached by that number, the interviewer was prompted by the CATI system to ask if the contact person knew the sampled person and, if so, if he or she knew a phone number where the person could be reached.

    a. If this request was successful, the interviewer entered the new phone number into the CATI system. The system then determined if the number was a completely new number or was identical to one already in the system. If new, the system added it to the priority list and it immediately became the priority number for future contact attempts.

    b. If this request was not successful, the system determined if there was still an untried number in the priority list and, if so, made the next one the number for future contact attempts.

2. If, after repeated dialings to the currently prioritized number, there was no resolution as to whether the sampled member could be reached at that number, the system determined if there was still an untried number in the priority list and, if so, made the next one the number for future contact attempts.

There was no limit on the number of times this process of requesting a new phone number from a live contact at an unsuccessful number could be repeated. The CATI system called through every available number (from original sources or obtained during telephone interviewing operations) until the sample member was found or every available number had been called without successfully locating the member.

Further, only the actual sampled member was eligible to respond, not another family member, even if that person was a Blue Cross member; therefore, the interviewers confirmed that they had actually reached the specific sampled individual, and not, for example, someone else with the same or similar name, such as a child or parent of the sampled member.

3.6.1.4 Calling Rules Determined by Special Case Eligibility Rules

During data collection, two issues emerged that required the MATS 2007 design team to develop handling rules that were consistent with the study's research

objectives, statistical sampling practices, and data collection operational procedures. Their common feature was temporary residence out of state.

3.6.1.4.1 Snowbirds

Discussions among the survey sponsors and Westat produced a protocol for handling "snowbirds," residents who leave the state for warmer climates in the winter. MATS 2007 considered them as valid Minnesota residents on an extended vacation. They were, therefore, eligible members of both the RDD and Blue Cross samples, and MATS 2007 sought to interview them. However, because some MATS 2007 research questions address social and environmental factors and Minnesota policies and programs, the survey designers concluded that it was best to interview people only when they were physically within the state borders. If snowbirds currently dwelling out of state were identified during initial calls, arrangements were made to interview them upon their return to Minnesota, if they returned before the end of the data collection period. If they were not in the state during this period, they were not followed to their winter residence for an interview. Since they remained as eligible members of the sample, the latter group was classified as a form of survey non-response, rather than as ineligibles who could be dropped from the sample for response rate calculations and sample weighting.

3.6.1.4.2 College Students Living Away from Home (RDD and Blue Cross Samples)

A common protocol for RDD surveys (based on fairly complex rules that the U.S. Census uses to define household membership) is that students who reside away from home are considered members of the household if they live in group quarters (such as a dormitory) but not if they reside in private or small common residential units (generally defined as those occupied by nine or fewer unrelated individuals).

The research issues that arose in regard to snowbirds similarly applied to the out-of-home student situation, and the MATS 2007 designers slightly modified the commonly applied RDD rules for students.

Young adults who were found to be attending college in Minnesota were eligible to be interviewed, even if they were not currently residing in the household at the

phone number of record (most likely their parents' house). The procedure was to call them at any phone number where they could be reached to conduct the interview, including calling back at the number of record if they would be available there before the end of data collection.

Young adults who were found to be attending college outside of Minnesota were classified as not eligible for the study, because they would be less exposed to the social and environmental factors and Minnesota policies and programs that were a major focus of MATS 2007. They were dropped from the sample for response rate calculations and sample weighting.

The RDD screening protocol and the Blue Cross contact protocol did not seek to explicitly identify students. The protocol was applied only in the situation where a student's status as residing temporarily away from home happened to emerge in conversation, whether with the sampled student or with someone else in the household. If this happened, the interviewer was instructed to ask if the student was attending school in state or out of state. If out of state, the interviewer recorded the case as a special problem with detailed comments for review by supervisory staff. If the supervisor concurred with the determination, the case was coded as ineligible; if not, it was reactivated with instructions to continue to pursue the interview with the student wherever he or she could be reached.

3.6.2 Home-based Interviewers

As noted above, MATS 2007 used traditional call center-based telephone interviewers and, mainly, telephone interviewers working from their homes. Distributed call center operations using home-based telephone staff networked though advanced web-based platforms are a growing trend in the survey research, customer support, and telephone counseling professions (such as tobacco telephone quitlines).

Regardless of their location, the MATS 2007 center-based and home-based interviewers accessed the same CATI system with a common CATI instrument, survey database, sample management system, call scheduler, and autodialer, employing a secure, web-based interface. All interviewers received cases from a common sample management database, and the call scheduling/case management

system transparently assigned cases to available interviewers. Using the web-based interface, all data collected by the home-based interviewers were entered in real time into the same central survey database accessed by the center-based interviewers, maintained within Westat's physical facility and behind Westat's software firewall. Sample identifying information, questionnaire text, and response data were only visible on the interviewers' screens; no data could be copied or saved electronically or printed locally.

Home-based interviewers were likewise monitored and received feedback using the same protocol, methods, and interface as for center-based interviewers. See section 3.7.1 for more information about interviewer monitoring.

3.6.3 Refusal Conversion

In scientific surveys, it is standard practice to recontact people who initially refuse to participate in an interview, in a second attempt to persuade them to participate. This refusal conversion process is designed to increase the sample size and response rates, and also to reduce bias associated with including in the sample only those who are most inclined to respond, i.e., those who respond immediately to the participation request. This section describes the MATS 2007 refusal conversion process for the RDD screeners, the RDD interviews, and the Blue Cross interviews. See Section 4.3.3. for the quantitative results of these refusal conversion efforts.

When interviewers encountered reluctant respondents, they first attempted to avoid the refusal by addressing any concerns expressed. When that attempt was unsuccessful, the interviewer coded the case a refusal and completed a non-interview report form. Included in this form was a brief description of the reason for and the strength of the refusal. The strength of the refusal was coded mild, firm or hostile (adamant). Mild and firm refusals were mostly determined at the discretion of the interviewer. Typically refusals were designated mild if the respondent hung up without explanation and firm if some type of reasoned explanation was provided. The hostile designation was reserved for respondents who used vulgar language or were threatening.

Each refusal case was withheld from additional call attempts for a cooling off period of at least 13 days. After this cooling off period, specially trained

interviewers attempted to recontact the household (for screener refusals) or the selected adult (for the extended interview), to persuade them to participate in the survey. If the respondent refused a second time, the case was finalized as a refusal and no further attempts were made to contact them. The interviewers selected for these conversion attempts demonstrated an above-average comprehension of the study objectives and ability to share this knowledge with the respondent. They received additional training to aid them in dealing with challenging situations.

3.7 Data Quality Assurance

Data quality assurance for MATS 2007 took a variety of forms prior to, during, and following data collection. Prior to data collection, data quality assurance was addressed through the questionnaire design, specification, and testing process described elsewhere. The valid generalizability of the collected data to the overall population is further assured by a well-designed and scientifically drawn sample. The techniques for designing and drawing the sample are described throughout chapter 2; diagnostic quality control checks prior to data collection are specifically described in Section 2.3.

The present section addresses measures implemented to assure the quality of the data as collected during and following data collection. Such measures include monitoring interviewers and providing feedback to them, conducting ongoing review of the actual data captured in the CATI system during data collection, ongoing monitoring of sample performance during data collection, and in the processing of the data into the final database once data collection is ended.

3.7.1 Interviewer Monitoring

Westat uses a silent monitoring system that allows supervisors to listen on the phone and to watch interviews on the CATI screen in real time without the interviewer or respondent knowing that they are being observed. Monitoring reports are completed for each monitoring session and reviewed with the interviewer during each shift. This provided the opportunity to reinforce good skills and coach interviewers in areas needing improvement in a timely manner. Approximately 10% of all interview time was monitored.

3.7.2 Data Cleaning and Editing During Data Collection

The primary method for assuring the quality of the collected data is to address this objective, before the fact, in the design and programming of the CATI questionnaire, in the data collection protocol developed, and in the training of the interviewers in general best practices and the specifics of the MATS 2007 questionnaire and interviewing protocol. All of chapter 3 up to this point has addressed these issues in detail.

Even though the CATI system controls all skip patterns and allows only valid ranges of values to be entered by the interviewers, Westat data managers conducted two additional complete and independent reviews of the collected data after the fact.

The first review consisted of continuous review of the frequency distributions of every survey variable during the data collection process. The CATI data manager reviewed the frequencies for each variable on a regular basis to check for any inconsistencies in the skip patterns or range violations. While rare in a well-tested CATI system, such errors may occur because of unusual situations not anticipated in the design or not revealed during testing. Such continuous quality assurance allows discovering any such problems early in the process, making necessary corrections, and recalling affected respondents to obtain corrected data (data retrieval). This process revealed one error in the CATI questionnaire programming for MATS 2007. The nature of the error and data retrieval activity is described further in Section 3.7.2.1.

The CATI data manager also reviewed comments noted by the interviewers in the CATI system. These comments might have been notes made by the interviewers themselves, or might have been extraneous comments by the respondents and recorded by the interviewers. Often the comments required no further action. In some instances, the comment could be an indication that the respondent corrected their answer to a previous question, or the interviewer was unsure how to code a particular response. In these cases, the CATI data manager made any necessary edits to the data or referred the case to a MATS data collection manager for a decision. Any such edits were documented in the data edit log, which contains

both the original value that was recorded in the interview as well as the new, updated value for each respective variable where an edit occurred, along with a brief description detailing the reason for the edit.

3.7.2.1 Detected CATI Programming Error and Data Retrieval

Cases with a smoking status of 'F1' (former smokers who had not smoked within the past 30 days) and a response to question D12 (About how long has it been since you last smoked cigarettes regularly?) of "less then a year ago" were to be asked question E1. Due to a programming oversight, if the respondent answered D12 using the unit of days, the instrument incorrectly skipped them past E1, and, based on the resulting absence of a value in E1, also skipped E2.

When the CATI data manager noticed the incorrect pattern, she alerted the survey operations manager and CATI programmer of the anomaly. The CATI programmer quickly determined the source of the problem in the CATI code, corrected it, tested the correction, and moved the corrected version into production.

Prior to the correction, a total of 16 cases were affected, 15 in RDD and 1 in the Blue Cross sample. Westat developed a hard copy contact form and 2-item questionnaire. The form contained the necessary contact information, an explanatory script and request for assistance in answering the missing questions, and the question text and response recording blocks. Westat assigned one dedicated interviewer to conduct the data retrieval, who succeeded in recontacting and collecting the data from 14 of the 16 respondents. The data from the hard copy retrieval forms were keyed into the CATI database by the CATI data manager.

3.7.3 Data Cleaning and Editing Following Data Collection

After data collection was completed, the MATS 2007 data delivery manager developed a SAS program that independently tested the data integrity rules and ascertained the follow through of all skip patterns. This SAS program served as a second layer of quality control to ensure the accuracy of the data integrity rules specified for the instrument. The program produced a detailed, case-by-case, variable-by-variable report if any errors were encountered. Errors in this context refers to instances where data was either missing, or data was present where it should not have been. The data manager reviewed the error report and made any

necessary corrections to the data to accurately satisfy the data integrity rules of the instrument. The data delivery manager also rechecked each variable for values outside of the allowed ranges. All edits made to the dataset were documented in the same data edit log that was used for edits made by the CATI data manager during data collection.

3.7.3.1 Post-coding of Verbatim Text in 'Other-Specify' Questions

Once the data were cleaned, additional processing of the data occurred as a result of reviewing and recoding the text responses to the various open-ended 'Other-Specify' questions that appear in the MATS 2007 questionnaire. ClearWay, Blue Cross, and Westat survey managers collaborated closely on all such post-coding and recoding decisions. In addition to creating post-codes for the 'Other-Specify' responses, this process also identified some instances where a categorical response to an earlier question in the same sequence as a given 'Other-Specify' question should be recoded, based on the additional information that the respondent furnished in the 'Other-Specify' response.

All updates made to the data as a result of this process were stored in the final dataset in newly created variables, and the data as originally collected were preserved in historical variables in the dataset. Westat documented the process in a review and recoding protocol document; a database at the record and variable level that recorded all 'Other-Specify' text responses, post-codes, and recodes; and a set of summary tables of the outcomes of this process on each affected survey variable.

3.7.4 Sample Performance Monitoring During Data Collection

Throughout the data collection process, it was vital to monitor several outcomes of the data collection process, mainly to project estimated final totals from interim results and determine what adjustments were needed or possible, in order to support achieving the targeted number of completed interviews for the various sample groups. Aside from standard weekly reports of case dispositions for the RDD and Blue Cross samples, there were four areas of particular interest to the survey sponsors and Westat technical managers.

1. From a sample management perspective, it was necessary to monitor the overall yield of completed interviews resulting from the sample release

groups activated at the outset and then in subsequent releases, to assure achieving the target number of completed RDD and Blue Cross interviews.

2. Since MATS 2007 oversampled African Americans, it was desirable to monitor the yield of completed interviews with African Americans.

3. Since MATS 2007 oversampled young adults, it was desirable to monitor the yield of completed interviews with young adults.

4. Since overall smoking prevalence was the primary measure to be estimated by MATS 2007, it was desirable to monitor this during data collection, to assess whether there was an unreasonably large divergence from the trend or estimates produced by other surveys, such as BRFSS. If such divergence was observed, investigating whether it was being artificially affected by some aspect of the survey design would be in order, so as to make corrective adjustments early in the data collection process to compensate.

3.7.4.1 Monitoring the Overall Number of RDD and Blue Cross Interviews

The yields of completed RDD interviews were slightly lower than initially projected during the sample design process. Examination of the monitoring reports revealed that this was due to two principal factors, a lower than expected rate of residential numbers in the sample of random phone numbers and a lower than expected response rate. The absolute yields of completed Blue Cross interviews were higher than projected in the design process. This higher yield did not result from the cooperation rate, however, which was actually lower than expected.[5] In fact, cooperation is only one factor affecting sample yields. Other factors discussed below offset this lower than anticipated cooperation rate. It is standard practice in survey research to initially release fewer of the originally projected number of sample release groups, then monitor sample performance to determine if more sample is needed. In response to each of these trends, the statisticians and operations managers updated the projections at several points during data collection and released additional RDD sample and refrained from releasing any

[5] The cooperation rate is the number of those who agreed to participate in the survey as a percentage of all those with whom it was possible to make contact and request their participation.

additional Blue Cross sample. Chapter 4 contains additional details about sample sizes.

Since projections were based on residency rates experienced in national RDD samples, the conclusion was that the incidence of working residential numbers across all area code/exchange combinations within Minnesota was lower than the national average. The ratio of such combinations to total active residential numbers will vary by state, based on the historical creation of area codes and exchanges, telephone central office capacity and technology issues, patterns in the cancellation of services, telephone company capacity projections, and so forth.

For the RDD sample, extended interview cooperation rates were close to expectation. However, the screener cooperation rates for the RDD sample were lower than expected, particularly in the African American high density stratum. As discussed in detail in the following section, this was a contributing factor to the lower than expected African American yield. While the cooperation rate associated with the Blue Cross sample was also lower than expected, two factors more than offset this lower than expected cooperation rate. First, there were fewer missing phone numbers on the Blue Cross frame than allowed for in the sample design projections. Second, the telephone number vendor look-up process yielded a higher than expected proportion of telephone numbers for those members who did not have a phone number on the sampling frame. The result of these two factors was that MATS 2007 achieved higher than expected success in locating the Blue Cross members by phone. The net result was that less sample was required to achieve the goal of 5,000 completed interviews than what was originally designed.

3.7.4.2 Monitoring the Number of African American Interviews

The total number of African American interviews would be the sum of the RDD and Blue Cross interviews with respondents who self-reported their race as African American during the extended interview. There was no direct method to sample African Americans in the RDD sample or the Blue Cross sample, but the RDD sample design used the indirect method of oversampling telephone exchanges covering census blocks with higher proportions of African Americans. The initially projected number of African American RDD interviews was based on assumptions

that the actual distribution of African Americans found through the random numbers selected in the low-density and high-density exchanges would be approximately the same as the overall distribution in the census blocks covered by the these two groups of exchanges. It was also based on the assumption that household screener and extended interview response rates among African Americans would be similar to those for the rest of the population.

The MATS 2007 sample design initially targeted 450 African American interviews from the RDD sample, plus whatever small increment would be encountered among the Blue Cross respondents. Given the smaller Blue Cross sample size and the expectation that African Americans would be no more prevalent in the Blue Cross sample than in the Minnesota population at large, the amplified RDD African American sample was the main tool for achieving larger numbers of African American respondents and, hence, the main focus for monitoring African American interview yields during data collection.

By the end of April 2007, it was clear that there would be a shortfall from the hoped-for 450 RDD African American interviews, based on yields to date and projections from currently active and prospective RDD release groups. Since race was determined only at the very end of extended interview, it was difficult to precisely determine the source of this shortfall, which may have been due to lower screener or individual response among African Americans than originally assumed or to lower prevalence of households with African Americans among the sampled RDD numbers than indicated by the U.S. Census statistics for Minnesota. Hence, there was no precise way of knowing if households with African American membership were refusing the screener or if sampled African American individuals were refusing the extended interview at a higher rate than others; similarly, there was no way of knowing if other forms or non-response were more prevalent among African Americans.

An indirect, imprecise measure of the response rate among African Americans was the higher screener refusal rate and the higher extended interview refusal rate in the high-density African American RDD sample strata, compared to the low-density strata. While there was no way of knowing if the refusal rate was higher among the African Americans in the high-density strata (and it is important to keep

in mind that they still represented a minority percentage in most of the high-density strata 100-blocks), this statistic was at least suggestive of the possibility of a differential response rate negatively impacting the yield of African American respondents.

Westat explored several options to increase the newly projected yield of African American respondents while holding constant the total number of RDD interviews at 7,500. These were:

1. Implement 100% screener refusal conversion in the high-density stratum for all existing and future screener refusals;

2. For future RDD sample releases needed to achieve 7,500 completed interviews, include all of the numbers in the high-density strata and whatever additional numbers would be needed from the low-density stratum; and

3. Create a new super-high density stratum, defined as those RDD numbers in census blocks with 25+% African American households, release all such numbers in the remaining release groups.

By itself, Option 3 would still have fallen far short of the original target for African Americans and it would not have yielded enough total interviews to reach the overall RDD target of 7,500. If adopted, it would have necessitated acquiring additional RDD sample numbers drawn from a 25+% stratum.

It is important to realize that, since it was impossible under the MATS 2007 design to subsample on race at the point of the household screener interview, a given quantum of released RDD numbers would produce mainly non-African American interviews even in the higher-density strata. Whatever revision of the sample design was implemented would have had two constraints: the overall target of 7,500 and the African American target of ~450 RDD interviews. A further constraint was the relative, incrementally higher, cost of sampling from higher-density strata, since these had been found to have lower household prevalences and lower response rates than the other strata. Another factor was the effect on overall precision of increasingly more disproportionate weights in the overall sample that would result from additional oversampling. Yet another factor was the trade-off of

the additional cost and additional statistical precision for additional African American sample. Any given modification of the sampling design for the African Americans would result in some additional cost. That same modification would also result in some gain in statistical precision of the survey estimates for the African American population. For any sample design, one issue is the ratio of the cost of implementing it to the precision achieved. For example, at a certain point, small additional increases in precision can require large increases in sample size, making them not cost-effective to achieve. For the MATS 2007 African American oversample, the modification of the sample design needed to strike a balance between the cost of the additional sample and the gain in precision to achieved from it. Finally, the effect of adopting any new option had to be assessed in terms of how the several factors determining sample yields in the original design and from the new options would interact, in terms of achieving the overall sample size of 7,500 completed RDD interviews; these included the different overall household prevalences and response rates encountered in different strata; the possible differences in African American household prevalences and response rates, across strata and also compared to non-African Americans within strata; and the need to accept all respondents yielded by each release group regardless of stratum and racial distribution.

Based on these considerations, Westat prepared a number of detailed analyses of sample yields, precision estimates, and estimated costs for the three options. Options 1 and 2, even in combination, would produce relatively few additional African American interviews and would have no appreciable impact on the reliability of estimates from the African American sample. Implementing Option 3 was projected to produce approximately 300 African American interviews, after adding together those already completed, those projected from cases still being worked, and those to be realized through Option 3.

While the 300 interviews still fell short of the 450 originally planned, it represented a balance among cost, relative gain in precision of the African American survey estimates, minimizing further reductions in the precision of the overall estimates resulting from the additional oversampling of a subpopulation, and maintaining the original sample size of 7,500 RDD interviews. The original target of 450 African

Americans was driven less by the absolute level of precision that a sample of that size could provide; rather, it was driven more by expectations of what could be achieved at reasonable cost and with reasonable effect on the overall survey precision, given the effect that greater or lesser oversampling of African Americans and young adults in various proportions would have on overall precision. The MATS Advisory Panel selected Option 3, and Westat acquired and released RDD sample from the super-high density stratum. When data collection concluded, the RDD sample yielded a total of 290 African American respondents. The Blue Cross sample included another 45 African Americans, for a total of 335 in the combined sample.[6]

3.7.4.3 Monitoring the Number of Young Adult Interviews

As with the number of African American interviews, the total number of young adult interviews was planned as the sum of the RDD and Blue Cross interviews with young adults. Two oversampling procedures were used to increase the number of young adult interviews. The first was applied to the RDD sample. This procedure oversampled young adults in households where both young and older adults resided. In such households, young adults were sampled at four times the rate of older adults. The second procedure was applied to the Blue Cross sample. Since age was on the Blue Cross membership record, young adults could be directly oversampled. The oversampling rate applied to young adults on the Blue Cross frame was three times that of older adults and was applied across all relevant plan types.

It became clear early in the data collection process that the target for the young adult yield from the RDD sample would not be met. Many possible sources of this shortfall were identified, including a lower response rate for young adults, coverage problems associated with young adults in RDD surveys in general, and problems classifying young adults and older adults in mixed-age households (households where both young and older adults both reside). In theory, the first source (a low response rate) could be somewhat countered by further increasing the

[6] African Americans were identified during data collection based on their responses to MATS 2007 questionnaire item J4b, Are you Black or African American. For weighting and analysis, various other definitions were employed, based on combinations of multiple racial heritages, Hispanic ethnicity, and country of birth.

oversampling rate of young adults in mixed-age households. However, it any increase in the already high oversampling rate would have a negligible effect in increasing the number of sampled young adults and therefore was not implemented. The second source, coverage issues associated with young adults in RDD surveys, can be addressed only in the design of the study.[7]

The final source of the lower young adult RDD yield was identified during MATS 2007 data collection. The problem can be described as follows. Each adult identified in a multi-adult household during the screening interview was classified by the screener respondent as either an adult 18-24 years old or an adult 25 or older, in response to an explicit screening question. However, analysis of the screener and interview data revealed that, in more instances than could be explained by simple respondent error, the sampled adult sometimes reported his or her age during the actual interview differently from the age group reported by the screener respondent. For example, if a sampled person was reported as a young adult during the screening interview, he or she would self-report as an older adult during the actual interview. When it became apparent that the young adult yield was lower than expected, Westat diagnosed the problem by cross-tabulating adult age status in the screening interview with adult age status in the extended interview. These tables revealed that the discrepancy was much more common when a young adult was identified and sampled during the screening interview and was very rare when an older adult was identified and sampled. Westat's review of the tables and the screener questionnaire determined that the screener wording surrounding the adult age classification and the request for an interview with the sampled person who fell into the desired age category was potentially unclear to some screener respondents. Westat proposed slight modifications of the relevant screener language. Once these modifications were finalized and accepted by the MATS Advisory Panel, they were immediately implemented in the CATI system. Review of the diagnostic cross-tabulations revealed that, once the wording change was in place, the age

[7] Future MATS may need to consider alternative designs to the traditional RDD method used for the first three cycles. For example, a cell phone component may need to be included in the next cycle of MATS to supplement the RDD component young adult yield since a sizeable proportion of young adults have moved to cell-only or cell-mostly households (Blumberg & Luke, 2008). It is important to keep in mind that this is only an issue if the smoking rate among young adults in cell-only or cell-mostly households is different than the smoking rate of young adults in landline households.

misclassification virtually disappeared and the yield of young adults improved. However, this small clarification could not compensate for the response rate and coverage issues endemic to young adults in RDD surveys.

Other than the screener wording change, no overt adjustments were made to try to increase the yield of young adults in the RDD sample. Further, the yield of young adults in the Blue Cross sample was actually higher than expected and offset somewhat the shortfall on the RDD side. The final, combined sample included 1,469 young adults.

3.7.4.4 Monitoring Smoking Prevalence Rates

Smoking prevalence rates were monitored throughout data collection, and particularly at the point when the completed sample size was large enough to make the interim calculated rate predictive of the final estimates. The reported prevalence rate from the MATS 2003 survey was 18 percent. The expectation was that this percentage would drop slightly in MATS 2007, but a large difference between the MATS 2003 and the interim MATS 2007 prevalence rates in either direction could indicate a potential problem with the questionnaire, the interviewing, or the data collection procedures that manifested itself in regards to smoking status classification.

Initial unweighted smoking prevalence rates raised concern because they were lower than what was felt to be plausible. Further investigation revealed that, as expected, the older population in Minnesota were more likely to respond to the survey. It is well established that older people have a lower smoking prevalence than younger ones. Westat quickly developed a rough weighted estimate of the smoking prevalence rate using the completed MATS 2007 sample to date. The rough weighting process accounted for differences in the observed age distributions of respondents and the true distributions in the Minnesota population as reported by the U.S. Census. Applying these rough weights brought the interim estimate of smoking prevalence up closer to, but still below, the MATS 2003 estimate. This interim estimate provided assurance regarding the classification of smoking status during the administration of the MATS 2007 survey. Monitoring of the

unweighted and rough-weighted smoking prevalence estimate continued until the end of data collection, with no further cause for concern.

4. Data Collection Results

Chapter 4 presents various statistics summarizing the outcomes of the MATS 2007 data collection, separately for the RDD and Blue Cross samples. The key statistics presented are the call dispositions and the response rates for these two sample groups.

4.1 Completed interviews

Table 4-1 presents the overall number of completed interviews for the RDD and Blue Cross samples, with breakouts of the young adult and African American respondents. The young adult and African American categories are not mutually exclusive.

Table 4-1. MATS 2007 Completed Interviews, by Sample Type

Completed interviews	RDD	Blue Cross	Total
Overall	7,532	5,048	12,580
Young adults	475	994	1,469
African Americans	290	45	335

4.2 Telephone Interviewing Results

4.2.1 RDD Sample Calling Dispositions

Tables 4-2 and 4-3 show the detailed dispositions for all of the sampled RDD telephone numbers that were released to data collection. Table 4-1 shows the dispositions for the screening of the 94,467 sampled RDD numbers; Table 4-2 shows the dispositions of the 10,025 numbers for which a household screener was completed and from which an adult was selected for the MATS 2007 interview. These tables tabulate the actual disposition categories that Westat employed to manage the sample for the MATS 2007 interviewing operations. They also show the standard AAPOR disposition codes to which each lower-level MATS 2007 disposition category is mapped. (AAPOR is the American Association for Public Opinion Research, to which many survey researchers belong and whose members

have established various standards for scientific survey research that are widely accepted.)

Note that AAPOR dispositions account for each sampled phone number. Operationally, Westat accounted for the results of the sampled household members in the 10,025 completed household screeners as a second level of detailed disposition codes. For this reason, there is no corresponding AAPOR code for these cases in Table 4-2; rather, the AAPOR codes for these cases appear in Table 4-3. Separating the cases into the two operational stages provides a clearer understanding of the RDD data collection outcomes, while still allowing all 94,467 sampled telephone numbers to be classified as to their outcomes according to the standard AAPOR disposition codes.

4.2.2 Blue Cross Sample Calling Dispositions

Table 4-4 shows the detailed dispositions for all of the sampled Blue Cross members that were released to data collection. As discussed in Section 4-3, AAPOR standard disposition codes are designed for use in RDD surveys and do not apply telephone list samples, like the Blue Cross sample; consequently, Table 4-4 shows only the Westat disposition categories.

4.3 MATS 2007 Response Rates

This section presents the survey response rates for MATS 2007 RDD and Blue Cross samples. High response rates are desirable because, all other things being equal and absent any better measures, a higher response rate gives reason to believe the responding sample is a better representation of the study population than it would be if the response rate were lower. There are numerous rules and formulas used to calculate response rates, and there are different types of response rates. The method chosen for calculating response rates should offer the best, most defensible approach to assessing the representativeness of the responding sample, compared to the population it is meant to represent.

Table 4-2. MATS 2007 RDD telephone number sample dispositions

AAPOR code	Description	Count	Percent
N/A	Completed Screener	10,025	10.6%
4.7	There is no one in the household age 18 or older to do the screener.	19	0.0%
4.7	The sampled telephone number rings into a household not located in Minnesota	1	0.0%
2.331	Language Problem: unable to communicate due to a hearing or speech problem or unable to reach an English speaking household member.	560	0.6%
2.2	The maximum calling algorithm has been fulfilled. At least one "human" contact has been made at the number and there are no refusals or language problems in the call history for the household.	563	0.6%
3.13	The calling algorithm has been fulfilled with no "human" or answering machine contact.	3,339	3.5%
4.5	The telephone number was identified as non-residential during business purge preprocessing prior to CATI load.	6,329	6.7%
3.14	The maximum calling algorithm has been fulfilled for a telephone number and only answering machine contact was made.	2,682	2.8%
2.35	Non-response: other. Case for which no other final result code is applicable.	2	0.0%
4.5	The telephone number called was not a residential number. Included are businesses, institutions, agencies, modems, public facilities, vacation homes, group quarters.	3,703	3.9%
4.3	The telephone number was identified as non-working during Tri-tone match processing prior to start of calling.	49,841	52.8%
4.3	The telephone number was found to be not working when called.	9,444	10.0%
2.111	Refusal – Household screener respondent refused to be interviewed or broke off during the screener interview.	4,027	4.3%
2.111	Refusal – Screener refusal not flagged for refusal conversion.	3,932	4.2%
	Total	**94,467**	**100.0**%

Table 4-3. MATS 2007 RDD extended interview sample dispositions

AAPOR code	Description	Count	Percent
1.1	Completed interview.	7,532	75.1%
4.1	Subject does not live in the state of Minnesota.	8	0.1%
2.332	Language Problem: unable to communicate due to a hearing or speech problem or the selected respondent was unable to speak English.	71	0.7%
2.2	The maximum calling algorithm has been fulfilled. At least one "human" contact has been made at the number.	621	6.2%
2.31	Non-Response: subject deceased.	10	0.1%
2.3	Non-Response: interviewer error – sampled respondent not interviewed.	6	0.1%
2.22	The maximum calling algorithm has been fulfilled for a telephone number and only answering machine contact was made.	1	0.0%
2.35	Other Non-Response	3	0.0%
2.21	Selected respondent not available in field period	4	0.0%
4.5	The telephone number called was not a residential number. Included are businesses, institutions, agencies, modems, public facilities, vacation homes, group quarters.	9	0.1%
2.32	Non-Response: subject physically or mentally incapable of completing interview	25	0.2%
4.3	The telephone number was found to be not working when called.	51	0.5%
4.1	Enumeration error – The respondent enumerated in the screener and selected for the extended interview is not a member of the household (typically occurs when visitors or family members living away are erroneously reported as household members.)	115	1.1%
4.1	Other out of scope – The case is out of scope and no other final code applies.	1	0.0%
2.112	Refusal – The selected respondent or a gatekeeper refused the interview or the selected respondent broke off during the interview and refused to continue.	1,567	15.6%
2.112	Refusal – Results from a call to the Minnesota Department of Health.	1	0.0%
	Total	**10,025**	**100.0%**

Table 4-4. MATS 2007 Blue Cross sample dispositions

Description	Count	Percent
Completed interview.	5,048	44.8%
Subject does not live in the state of Minnesota.	31	0.3%
Language Problem: unable to communicate due to a hearing or speech problem or sampled person was unable to speak English.	97	0.9%
The maximum calling algorithm has been fulfilled. At least one "human" contact has been made.	692	6.1%
Ineligible: Subject deceased.	21	0.2%
The sampled person was not located.	2,707	24.0%
The sampled person was not available in field period.	8	0.1%
Non-Response: subject physically or mentally incapable of completing interview	38	0.3%
Duplicate of phone number in RDD sample.	166	1.5%
Other out of scope – The case is out of scope and no other final code applies.	6	0.1%
Refusal – The selected respondent or a gatekeeper refused the interview or the selected respondent broke off during the interview and refused to continue.	2,278	20.0%
Refusal – Results of a call to Blue Cross.	177	1.6%
Total	**11,269**	**100.0%**

4.3.1 MATS 2007 RDD Response Rate

The MATS 2007 response rate is calculated using one of the standard AAPOR rates formulas. All of the AAPOR response rate formulas collapse the numeric AAPOR disposition codes to one decimal place and then assign them to the broad outcome categories in Table 4-5.

Table 4-5 AAPOR response rate formula disposition categories

Formula disposition category	Formula category meaning	Collapsed disposition code
I	Completed Interview	1.1
P	Partial Interview	1.2
R	Refusal and break-off	2.1
NC	Non-contact	2.2
O	Other	2.3
UH	Unknown if household occupied	3.1
UO	Unknown, other	3.2

In addition to the above disposition codes, the response rate formula uses the following symbol in the formula:

e = Estimated proportion of cases of unknown eligibility that are eligible

AAPOR has two similar response rate formulas that are relevant to MATS 2007, RR3 and RR4. The only difference between AAPOR RR3 and RR4 is that RR3 excludes partial completes from the numerator and RR4 includes them, resulting in a slightly higher response rate. Partial completes are those where the respondent stopped before the end of the interview, and it was never possible get him or her on the phone again to complete the interview, for whatever reason. Typically, these "breakoffs" occur because the respondent refused to continue and could not be converted, or the respondent was unable to continue (perhaps due to time constraints or an emergency). It is appropriate to include partial completes in calculating the response rate if two criteria are met. First, the partial completes should be included in the weighted analysis file. If they are not, then the additional "representativeness" they contribute to the response rate is not real. In effect, the study takes credit for the partial complete in the response rate, but then ignores the cases in analysis. Analysts often prefer to leave partial completes out of the analysis file because of the problems their missing data cause. This consideration is the basis of the second criterion for including partial completes. Typically, if partial completes are used at all, there is some standard established for sufficient data

being collected to include the partial complete. The definition of sufficiency varies: critical questions are defined in advance and partial completes are included if all or a certain percentage of critical questions are answered; a certain percentage of all questions must be answered; the respondent must have gotten past a certain point in the interview; and other such measures.

MATS 2007 did not include partial completes in the analysis file and did not assign a final sample weight to them. To be included in the weighted file used for analysis, an interview had to have reached the last question, J14, sex of respondent. Those that broke off before this point are accounted for in Table 4-3 among the breakoffs. Given these considerations, AAPOR RR3 (excluding partial completes from the numerator) is the proper formula to use for MATS 2007.[8] RR3 is given by:

$$\mathbf{RR3} = \frac{\mathbf{I}}{\mathbf{(I + P) + (R + NC + O) + e(UH + UO)}}$$

Response Rate 3 (RR3) estimates what proportion of cases of unknown eligibility is actually eligible. The proportion of such numbers to be retained in calculating the response rate is often referred to as "e," and the e-factor is applied to the phone numbers with unknown residential status after data collection has concluded. Since this occurs in the denominator of the response rate calculation, the lower the "e," the lower the denominator, and the higher the response rate. From AAPOR's Standard Definitions: "In estimating e, one must be guided by the best available scientific information on what share eligible cases make up among the unknown cases and one must not select a proportion in order to boost the response rate. The basis for the estimate must be explicitly stated and detailed. It may consist of separate estimates (Estimate 1, Estimate 2) for the sub-components of unknowns (3.10 and 3.20) and/or a range of estimators based of differing procedures. In each case, the basis of all estimates must be indicated."[9]

[8] MATS 2003 handled partial completes in weighting, file preparation, and analysis in the same way as MATS 2007, i.e., excluded them completely. However, MATS 2003 did include them in the numerator and denominator when calculating the RDD and Blue Cross survey response rates.

[9] The American Association for Public Opinion Research. 2008. Standard Definitions: *Final Dispositions of Case Codes and Outcome Rates for Surveys. 5th edition*. Lenexa, Kansas: AAPOR.

Westat's independent research has found that approximately 27% of numbers that are constant "No Answers" and 63% of numbers that are constant "Answering machines" are residential on a national level. For MATS 2007, Westat has calculated the RDD response rate using subfactor e's of 0.27 and 0.63 for "No Answers" and "Answering Machines," respectively. This produces a blended e of 0.443, weighted by the respective numbers of "No Answers" and "Answering Machines" in the sample at the end of data collection. While there is always a degree of uncertainty about the value of e, Westat believes this approach presents a better estimate of the response rate's representativeness than a uniform, single factor e.[10]

Another consideration is reporting an unweighted response rate versus a weighted response rate. For samples with unequal probabilities of selection, such as stratified samples and those with subsampling, weighted response rates are a better measure of the representativeness of the responders, for reasons similar to those why weighted data are more representative than unweighted: they account for the presence in the sample of individuals with certain characteristics (such as the propensity to respond) in proportions different from their natural distribution in the study population. This report presents weighted response rates for the RDD and Blue Cross samples.

The weighted version of the AAPOR RR3, $RR3_w$, is given by:

$$RR3_w = \frac{I_w}{(I_w + P_w) + (R_w + NC_w + O_w) + e(UH_w + UO_w)}$$

where the subscript w reflects the fact that the elements are the weighted totals of the cases falling into the different disposition categories.

[10] MATS 2003 used an e of 0.25. This is not an unreasonable number on its face, although the basis for choosing it was not stated. As noted, the lower the e, the higher response rate. The MATS 2007 e of 0.443 is higher than the MATS 2003 e of 0.25. All other things being equal, this will make the calculated MATS 2007 RDD response rate somewhat lower than if the MATS 2003 e were applied.

Table 4-6 shows the unweighted counts[11] of the cases that fell into each disposition category; it excludes the cases in Table 4-2 that were classified as non-working or non-residential numbers, since these are also excluded in the AAPOR RR3 formula. In calculating the weighted response rate, each case's base weight, adjusted for refusal conversion subsampling, was applied in the formula. The resulting response rate is 40.7%. Had the MATS 2003 e value of 0.25 been used instead of the MATS 2007 value of 0.443, the response rate would have been 42.7%.

Table 4-6. Unweighted case counts for RDD AAPOR Response Rate Formula 3

Description	AAPOR disposition code	Unweighted counts
Completed interview	I	7,532
Partial interview	P	0
Refusal and break-off	R	9,527
Subsampled for refusal conversion*	R_S	5,595
Not subsampled for refusal conversion*	R_{NS}	3,932
Non-contact	NC	1,189
Other	O	677
Unknown if household/occupied	UH	6,021
Ring no-answer*	UH_{NA}	3,339
Answering machine*	UH_{NM}	2,682
Unknown, other	UO	0

* The subscripted AAPOR disposition codes are not AAPOR notation; they have been created here to differentiate levels of these categories that are used in the weighted response rate formula, as described in the text.

It is important to keep in mind that this is the overall net response rate, across both the screening and extended interview stages. It also accounts for the fact that screener refusals were subsampled for conversion.[12]

[11] While it would have been arithmetically possible to calculate an unweighted MATS 2007 RDD response rate, this would not be meaningful, and would be misleading, due to such factors as the subsampling of screener refusals for conversion. MATS 2003 did not calculate a weighted response rate for the RDD sample.

[12] In calculating the weighted response rate, the weights of the cases not subsampled for conversion were set to zero, and their original weights were apportioned to the cases that were subsampled for conversion. Since some of the cases for which conversion was attempted yielded completed screener interviews and the rest did not, the overall response rate validity is preserved by this apportioning step.

4.3.2 Blue Cross Response Rate

AAPOR does not provide a formula for telephone list sample studies like the Blue Cross sample. RDD response rate formulas do not work for list samples, for the principal reason that every sampled case except ineligibles should go into the denominator of a list sample response rate. For a scientific random sample, it is not proper to exclude cases because they do not have a known phone number, cannot be located, cannot be contacted, or call the sponsor in response to an advance letter to refuse cooperation.

For the MATS 2007 Blue Cross sample, a case was considered ineligible[13] and dropped from the response rate calculation if the sampled member:

- No longer lived in Minnesota,
- Was a snowbird with a temporary but extended residency outside of the state through the end of the data collection period,
- Was found to be under 18 years old, or
- Was deceased.

In addition, as described in Chapter 2, Blue Cross sample members whose phone numbers were duplicates of phone numbers in the RDD sample were dropped as duplicate cases from the Blue Cross sample. They are also excluded from the Blue Cross response rate calculation.

Table 4-7 shows the unweighted counts of the cases that fell into the principal groupings of the Blue Cross disposition categories. The unweighted response rate is given by:

$$\textbf{List RR} = \frac{\textbf{I}}{\textbf{E}}$$

where E is the count of all eligible sampled Blue Cross cases.

[13] The Blue Cross sample frame was temporally defined by the membership list on January 1, 2007. It was not necessary for a Blue Cross respondent to still be a member on the date of the interview, and there was no screening of the Blue Cross sample for current Blue Cross membership during the interview.

In calculating the weighted response rate, each case's base weight, adjusted for refusal conversion subsampling, was applied in the following formula.

$$\textbf{List RR}_w = \frac{I_w}{E_w}$$

where E_w is the weighted count of all eligible sampled Blue Cross members.

As presented in Table 4-7, the unweighted Blue Cross response rate is 45.7%. It includes in the denominator all cases that should enter into the calculation. For comparison, the weighted response rate for the Blue Cross sample is 48.4%.[14]

Table 4-7. Blue Cross unweighted case counts and response rate

	Unweighted
Total members sampled	11,269
Members with phone numbers duplicated in RDD sample	(166)
Total non-duplicate members	11,103
Members determined ineligible	(31)
Members deceased	(21)
Total eligible members (E)	11,051
Total members not responding	(6,003)
Total members responding (completed interviews) (I)	5,048
Response rate *(Responded[I]/Eligible [E])*	45.7%

Table 4-8 presents the unweighted operational "completion rate" for the Blue Cross sample. The operational completion rate is a process indicator, reflecting only cases that were assigned to data collection. The cases that were not assigned are those that a) had a phone number that was a duplicate of a number sampled for RDD b) had no phone number on either the Blue Cross frame or from the phone number

[14] MATS 2003 also calculated a form of weighted response rate for the Blue Cross sample, by calculating within-stratum response rates then taking an average of these that was ratio-adjusted by the proportion of the overall Blue Cross frame that each stratum represented. However, this is a less exact form of weighting the response rate than applying the individual sample weight of each case in each disposition category, as in MATS 2007.

look-up process or c) contacted Blue Cross after receiving the advance letter to ask not to be contacted for the survey. This operational completion rate is not an indicator of the representativeness of the sample, but only of the success of the telephone interviewing operation in completing interviews with the cases assigned to it. The Blue Cross operational completion rate was 48.6%.

Table 4-8. Blue Cross sample operational completion rate

Total members sampled	11,269
Members with phone numbers duplicated in RDD sample	(166)
Members with no phone number	(493)
Members opting out	(177)
Total members assigned to data collection	10,433
Members determined ineligible	(31)
Members deceased	(21)
Total eligible members assigned to data collection (E)	10,381
Total members not responding	(5,333)
Total members responding (completed interviews) (I)	5,048
Operational completion rate *(Responded [I]/Eligible [E])*	48.6%

4.4 Selected Operational Statistics

This section presents some statistics that characterize various operational aspects of MATS 2007. The reported statistics are for cases that were assigned to data collection, with the exception of Section 4.4.2, which examines sources of non-response both for those cases that were actually assigned to data collection and for the overall Blue Cross sample, including those for which telephone contact could not be attempted.

4.4.1 Blue Cross Sample Stratum Yields

Unlike MATS 2003, MATS 2007 did not oversample any of the Blue Cross plan types and there were no targets for competed interviews by any of the plan types. Table 4-9, which is included only for reference purposes, shows the number of sampled records in each stratum that were assigned to data collection as discussed above, the number of completed interviews, and the unweighted percentage completed. This percentage is a completion rate, that is, a quantification of the sample yield in each stratum for those cases that were assigned to data collection.

Table 4-9. Attempted and completed Blue Cross sample cases by stratum

Stratum	Attempted sample records*	Completed interviews	Percent completed
MinnesotaCare, age 25+	224	110	49.1%
PMAP, age 25+	138	44	31.9%
Medicare Supplemental plans, age 25+	1,541	1,008	65.4%
Commercial plans, age 25+	5,681	2,843	50.0%
18-24 year olds (from all plan types)	2,849	1,043	36.6%
Total	10,433	5,048	48.4%

* Excludes members with no phone number, duplicates of RDD numbers, and members who opted out, but includes ineligibles and deceased.

4.4.2 Principal Sources of Non-response

4.4.2.1 Principal Sources of Non-response in RDD Sample

Table 4-10 summarizes the results for all RDD cases, after eliminating the known non-working and non-residential numbers. Consistent with the AAPOR RR3 formula, this table collapses the screening and extended interview processes into a single set of results. For example, a screener refusal in one case and a completed screener that resulted in a refusal of the extended interview in another case are treated identically and count as two refusals in this table. The largest source of RDD non-response was refusal to respond to the screener or extended interview: 9,527 (54.7%) out of the total 16,737 possible phone numbers that could have yielded a completed interview had a refusal at one of the two stages. Non-contact

Table 4-10. Primary sources of non-response in RDD sample, collapsed across screening and extended interview stages

	A. Including all screener & extended interview refusals		B. Including only subsampled screener refusals & all extended interview refusals	
	Count	% of non-respondents	Count	% of non-respondents
Total non-respondents	17,414	-	13,482	-
Refused	9,527	54.7%	5,595	41.5%
No contact	7,210	41.4%	7,210	53.5%
Both	16,737	96.1%	12,805	95.0%

was the second-largest source of RDD non-response, 7,210 cases (41.4%). Combined, these two outcomes accounted for 96.1% of the total non-response. These results are summarized in Table 4-10A.

There are two points to keep in mind in regard to the numbers cited in the previous paragraph. First, as explained in Section 4.3.1, some undetermined proportion of the non-contact cases at the screener level were not really households, and the response rate formula discounted a proportion of them as non-residential. The statistics presented in this section are purely operational and count all non-contact cases as non-response. Second, also as explained in Section 4.3.1, screener refusals were subsampled for conversion and a weighting adjustment accounted for this in calculating the response rate. The preceding paragraph ignored the subsampling and counted all refusals as non-response, including those for which no conversion was attempted. Table 4-10B omits the screener refusals that were not subsampled for conversion, which is a more suitable view of the operational sources of non-response. When these are omitted, the refusals to respond to the screener or extended interview amount to 5,595 (41.5%) out of the remaining total of 12,805 possible phone numbers that could have yielded an extended interview. Non-contact is unchanged, at 7,210 cases, representing the largest source of non-response at 53.5%. Combined, these two outcomes accounted for 95.0% of the total non-response.

It is also informative, and more straightforward, to look at sources of non-response for the RDD extended interview, that is, among the 10,025 adults selected for the interview from the completed screeners. Table 4-11 shows there were 9,901 eligible cases among the 10,025 selected. Most of those who were not eligible were individuals who, upon being contacted for the interview, were found to have been erroneously included as members of the household during the screener (e.g., guests, family members not currently residing in the household). There were 2,369 non-respondents, of whom the largest number were 1,568 refusals (66.1% of non-respondents and 15.8% of all eligible sample). Most of the remainder were individuals who could not be reached despite repeated attempts to do so, including the extra call attempts made beyond the BRFSS protocol parameters. These were 622 cases, or 26.3% of non-response and 6.3% of all eligible sample. Combined, refusals and maximum contacts accounted for 2,190 non-respondents, or 92.4% of all non response and 22.1% of all eligible adult sample.

Table 4-11. Primary sources of non-response in RDD extended interview sample

	Count	% of non-respondents (n = 2,369)	% of total eligible sampled (n = 9,901)
Total sampled	10,025	-	-
Ineligible/out of scope	(124)	-	-
Total eligible sampled	9,901	-	100.0%
Total complete	7,532	-	76.1%
Total non-respondents	2,369	100.0%	23.9%
Refused	1,568	66.1%	15.8%
Maximum contact attempts	622	26.3%	6.3%
Both	2,190	92.4%	22.1%

4.4.2.2 Principal Sources of Non-response in Blue Cross Sample

Table 4-12A shows that nearly all of the non-response in the Blue Cross sample came from two sources. Out of the 11,051 eligible sampled members (after removing RDD duplicates, ineligibles, and deceased), 6,003 were non-respondents. The largest source was the inability to locate the sampled member: 2,707 out of the

total 11,051 sampled members could not be located (45.1% of all non-respondents). Non-located cases (n=2,707) included 493 for whom no phone number was ever found (see Table 4-8) and the balance of 2,214 for whom at least one possible number was available but who were not found at any available number. The second-largest source of non-response was refusal to respond to the interview: 2,455 of the 11,103 sampled members refused the interview (40.9% of non-response). These refusals included 177 who refused by calling Blue Cross after receiving the advance letter and 2,278 who refused when contacted by an interviewer. Combined, non-locatables and refusals accounted for 86.0% of all non-response. Non-responders represented 54.3% of all eligible sampled Blue Cross members. Non-located individuals were 24.5% of the sampled eligibles and refusals were 22.2%. Combined, they are nearly half of all sampled members (46.7%).

Table 4-12A. Primary sources of non-response in the Blue Cross sample: all cases included in response rate calculation

	Count	% of non-respondents (n = 6,003)	% of total eligible sampled (n = 11,051)
Total sampled	11,269	-	-
RDD duplicate	(166)	-	-
Ineligible/deceased	(52)	-	-
Total eligible sample	11,051	-	100.0%
Total complete	5,048	-	45.7%
Total non-respondents	6,003	100.0%	54.3%
Non-located	2,707	45.1%	24.5%
Refused	2,455	40.9%	22.2%
Both	5,162	86.0%	46.7%

Table 4-12B is similar to Table 4-12A, except that it looks at non-response from interviewing operations, that is, it excludes the validly sampled cases for which data collection was not attempted because no phone number was found at all or the members opted out in advance. Among these cases fielded to interviewers, 2,214 still could not be located at the available phone numbers, or 41.5% of fielded non-

response. Another 2,278 members refused to the interviewers, or 42.7% of fielded non-response. Combined, non-locatables and refusals accounted for 84.2% of fielded non-response. Non-responders represented 51.4% of all fielded Blue Cross cases. Non-located individuals were 21.3% of the fielded cases and refusals were 21.9%. Combined, they are over forty percent of all fielded cases (43.2%).

Table 4-12B. Primary sources of non-response in the Blue Cross sample: cases included in operational completion rate calculation (fielded cases)

	Count	% of non-respondents (n = 5,333)	% of total eligible sample fielded (n = 10,381)
Total sampled	11,269	-	-
RDD duplicate	(166)	-	-
Ineligible/deceased	(52)	-	-
Total eligible sample	11,051	-	-
Not fielded (no phone number or opt-out)	(670)	-	-
Total eligible sample fielded	10,381	-	100.0%
Total complete	5,048	-	48.6%
Total non-respondents	5,333	100.0%	51.4%
Non-located	2,214	41.5%	21.3%
Refused	2,278	42.7%	21.9%
Both	4,492	84.2%	43.2%

4.4.3 Refusal Conversion Results

Recontacting people who initially refuse to participate in an interview is designed to increase the sample size and response rates, and also to reduce bias associated with including in the sample only those who are most inclined to respond. This section describes the results of the refusal conversion efforts for the RDD screeners, the RDD interviews, and the Blue Cross interviews, as summarized in Table 4-13.

Table 4-13. MATS 2007 refusal conversion

	RDD screener	RDD interview	Blue Cross interview
Assigned to data collection	38,297	10,025	10,433
Initially refused (#)	9,882	2,321	3,294
Ineligible for conversion	3,932	-	-
Eligible for conversion	5,950	-	-
Converted (#)	1,569	695	724
Converted (%)	26.4%	29.9%	22.0%
Total completed	10,025	7,532	5,048
Converted as percent of completed	15.7%	9.2%	14.3%

4.4.3.1 RDD Refusal Conversion Results

4.4.3.1.1 Screener Conversion

At the screener stage, the initial telephone contact resulted in a refusal to respond to the screener questions at 9,882 phone numbers. Of those initially refusing to complete the screener 5,950 were subsampled for a refusal conversion attempt, as previously described. The remaining 3,932 were not subject to conversion. After conversion attempts with the 5,950 initial refusals, 1,569 of these cases became completed screener interviews, representing a conversion rate of 26.4% and constituting 15.7% of the total 10,025 completed screeners.

4.4.3.1.2 Selected Household Member Conversion

Among the 10,025 household members selected for the MATS 2007 interview, 2,321 initially refused to respond to the interview. After conversion attempts, 695 of these completed the interview. This is a conversion rate of 29.9%, representing 9.2% of the total 7,532 completed RDD interviews.

4.4.3.2 Blue Cross Refusal Conversion

Among the 10,433 sampled Blue Cross members assigned to data collection, 3,294 initially refused to respond to the MATS 2007 interview. After conversion attempts,

724 of these completed the interview. This is a conversion rate of 22.0%, representing 14.3% of the total 5,048 completed Blue Cross interviews.

4.5 Interview Timing

Table 4-14 presents statistics about the length of time it took to administer the MATS 2007 interview, showing the mean, minimum, and maximum lengths for the overall sample, for the Blue Cross and RDD respondents, and for different smoking statuses.

Table 4-14. Interview lengths (minutes)

Sample Group	Mean	Minimum	Maximum
Blue Cross			
Current smokers	23.2	13.0	47.7
Former smokers	18.6	5.5	99.0
Experimenter smokers	15.0	4.8	52.4
Never smokers	14.3	4.2	38.7
All Blue Cross respondents	16.6	4.2	99.0
RDD			
Current smokers	24.5	6.6	50.6
Former smokers	19.4	5.0	72.2
Experimenter smokers	15.8	4.0	38.8
Never smokers	14.9	4.0	40.0
All RDD respondents	17.8	4.0	72.2
All respondents	17.3	4.0	99.0

Overall, the average interview lasted slightly over 17 minutes. The design targets were averages of no more than 25 minutes for current and former smokers and 15 minutes for never smokers. As seen in the table, the actual experience was very close or slightly under these targets.

As expected, interviews with current and former smokers took longer to complete than did those with experimental and never smokers, with the current smokers requiring the most time on average, owing to the largest number of questions

applying to their circumstance. However, the longest individual interviews were conducted with former smokers. The mean timings were similar for each smoking status across samples. The minimum timings were also consistent with the exception of the current smokers. With the exception of former smokers, no interview exceeded an hour in length.

5. Weighting

Sample weights are created so that population estimates can be made using the results of a survey from a sample of a finite population. For MATS 2007, three sets of weights were created: RDD-only weights, Blue Cross sample weights, and weights for a composite sample that combined the RDD and Blue Cross samples.

The RDD weights were created from the results of the RDD-only sample and were benchmarked to Minnesota population totals using the 2005 American Community Survey (ACS) population estimates as their source.

The Blue Cross sample weights were created so that estimates from the Blue Cross survey could be used to make inferences about the population making up the Blue Cross membership. Therefore, the Blue Cross weights were benchmarked to the Blue Cross sample frame counts.

A set of composite weights was also created. These composite weights resulted from combining the RDD and Blue Cross survey data into a single dataset for use in producing Minnesota statewide estimates. Following the creation of the composite weights, they were benchmarked to the ACS population estimates for the state of Minnesota, so that the combined data could be used to make inferences at the state level. By combining the results of the RDD and Blue Cross surveys and using the composite weights during analysis, reliability of survey estimates is increased compared to using the RDD sample alone, because of increased sample size.[15] In the sections that follow, the creation of these three sets of weights is described. Section 5.1 discusses the creation of RDD only weights, Section 5.2 describes the Blue Cross sample weights, and Section 5.3 the details the methods used to create the composite weights.

[15]Because of the complex sample designs and the merging of the two samples, the reliability of the estimates using the 12,580 interviews in the combined sample is less than would result from a simple random sample of the same size.

5.1 RDD Weights

The RDD weighting process started with the creation of household level weights, which are the inverse of their respective probabilities of selection. This is discussed in section 5.1.1. Subsequent adjustments made to these weights to account for sample deduplication; refusal conversion subsampling and multiple landline telephones in the household are discussed in Sections 5.1.2 thru 5.1.4. Person-level weights are discussed in Section 5.1.5, and the raking adjustments and raking dimensions are discussed in Section 5.1.6. Section 5.1.7 discusses the imputation procedure used to fill in missing data for variables needed when performing the raking adjustment.

5.1.1 Initial Household Weights

Each telephone number in the RDD sample was assigned an initial weight. This initial weight was computed as the inverse of the probability of selection of the telephone number. In MATS 2007, telephone numbers were drawn from a single frame of Minnesota telephone exchanges. The exchanges were split into three sampling strata defined by different levels of the proportion of African Americans living in those exchanges as determined by Census 2000 information. Table 5-1 shows the definition of the three strata. Initially, strata 2 and 3 were sampled at three times the rate as for stratum 1. Therefore, the initial weights for the oversampled strata were originally about one-third that of stratum 1. Well into data collection, additional sample was selected from stratum 3 in order to boost the African American yield. The additional oversample from this stratum reduced the stratum 3 baseweight further still.

Table 5-1. Definition of RDD phone number sampling strata for MATS 2007

Stratum	Definition
1 – Low density	Exchanges with less than 15% African Americans
2 – High density	Exchanges with between 15% and 25 % African Americans
3 – Super High density	Exchanges with at least 25% African Americans

Let $HHIW_{hi}$ denote the initial weight for the *i*-th sampled telephone number in the *h*-th stratum as defined in Table 5-1. Then the initial household weight is given by

$$HHIW_{hi} = \frac{N_h}{n_h},$$

where N_h is the number of telephone numbers in stratum *h*, and n_h is the number of telephone numbers sampled in stratum *h*.

Because the weights are adjusted within stratum, the subscript h will be omitted in the description of the weights in the following sections as a notational convenience.

5.1.2 Extra RDD Sample Deduplication Adjustment

Since the overall MATS 2007 sample design required sample from an RDD frame of telephone exchanges and sample of members from the Blue Cross membership frame, there was a potential for some telephone numbers on the RDD sample to appear as some members' phone numbers on the Blue Cross sample frame, a so-called duplicate number[16]. Prior to the start of data collection, if a telephone number was found in both samples, the member record containing it was removed from the Blue Cross sample. This removal of members from the Blue Cross sample was accounted for in the Blue Cross weighting process (see section 5.2 for details). After data collection started in both the RDD and Blue Cross samples, extra sample from the super-high density African American exchanges was drawn.

Similar to the duplication that existed between the original RDD sample and Blue Cross sample, there was a chance that additional duplicate records would be found between the supplemental RDD sample and Blue Cross sample. As described in Section 2.2.2., it is less biasing to drop Blue Cross members with duplicate numbers, rather than the duplicate RDD numbers, and that was the *a priori* procedure used when the RDD and Blue Cross samples were originally drawn. In the case of the

[16] As described above in Section 2.2.3, and explained in detail in memos to the Advisory Panel and in the MATS 2007 Comparability Report, RDD surveys use a sample of phone numbers and the Blue Cross survey sampled persons. For various reasons described in those documents, the presence of identical numbers on a phone number sample and on the administrative file records like the Blue Cross member list does not necessarily indicate that the same person has been sampled in both samples. However, to avoid potential burden, confusion, and likely duplication, the approved sample design specified the removal of the Blue Cross sample record when its telephone number matched an RDD sample number.

extra super-high density sample, however, any new duplicates had to be dropped from the extra RDD sample, since the previous RDD and Blue Cross samples were already released for data collection. In order to account for the numbers that were dropped from the extra sample as a result of also being found on the Blue Cross sample records, an additional adjustment was made to the RDD initial weights; however, it was not necessary to adjust for the duplicate numbers found in both the original and extra RDD samples, because the RDD initial weights account for the removal of these cases.

In order to make this adjustment for duplication, the weighting process first classified the RDD telephone numbers into one of three duplicate groups as shown in Table 5-2. The first group was comprised of the duplicate numbers in the extra sample. Recall that the entire extra sample was comprised of numbers in the super-high density stratum. The next group was made up of those duplicate numbers that were in the original sample and also in the super high density stratum. The last group was comprised of all other numbers not in groups one or two.

Table 5-2. Duplicate number groups

Duplicate group (DUPST)	Description
EXT_DUP	Numbers in extra sample and duplicates
ORG_DUP	Numbers in original sample and in super-high density stratum and duplicates
OTH	All other numbers, including numbers not in super-high density stratum and numbers in super-high density stratum but not duplicates

The RDD duplicate adjusted weight, $HHA1W_i$, is:

$$HHA1W_i = HHA1F_i * HHIW_i$$

where $HHA1F_i$ is the RDD duplicate adjustment factor computed as:

$$HHA1F_i = \begin{cases} \dfrac{\sum_{i \in EXT\_DUP, ORG\_DUP} HHIW_i}{\sum_{i \in ORG\_DUP} HHIW_i} & i \in ORG\_DUP \\ 0 & i \in EXT\_DUP \\ 1 & i \in OTH \end{cases}$$

where the groups *EXT_DUP*, *ORG_DUP* and *OTH* are defined in Table 5-1.

This approach has the effect of redistributing the sum of the base weights of all the *EXT_DUP* sample in the extra high-density sample to the *ORG_DUP* sample in the extra sample, thereby setting the *EXT_DUP* sample weights to zero in the final combined file, but retaining the original sum of the weights in the remaining RDD sample. The records that were given a weight of zero as a result of this adjustment were dropped from the weighting file and no further weighting adjustments were made to them. Additionally, the phone numbers that were designated as nonworking or business numbers were dropped from the weighting file at this point in the weighting process.

5.1.3 Refusal Conversion Adjustment

The base weights were adjusted to reflect the differential refusal conversion efforts made during data collection. Refusal conversion procedures were applied to a random subsample of household screener interview refusals. During sample selection, a flag was randomly assigned to approximately 60 percent of the telephone numbers. Refusal conversion at the household screener level was attempted only for the telephone numbers randomly assigned for conversion. In order to adjust the weights, telephone numbers were classified into screener refusal groups using the refusal status (if the respondent ever refused) and the value of the refusal conversion flag as shown in Table 5-3.

Table 5-3. Household screener refusal groups

Screener refusal group (REFST)	Respondent ever refused screener interview?	Refusal conversion flag	Description
NEVER_REF	No	N/A	Households where the screener respondent never refused
REF_CONV	Yes	1	Households where the screener respondent refused and refusal conversion procedures were applied
REF_NO_CONV	Yes	0	Households where screener respondent refused and refusal conversion procedures were not applied

The refusal conversion adjusted weight, $HHA2W_i$, is:

$$HHA2W_i = HHA2F_i * HHA1W_i$$

where $HHA2F_i$ is the refusal conversion adjustment factor computed as:

$$HHA2F_i = \begin{cases} \dfrac{\sum_{i \in REF\_CONV \ ,REF\_NO\_CONV} HHA1W_i}{\sum_{i \in REF\_CONV} HHA1W_i} & \text{If } i \in REF\_CONV \\ 0 & \text{If } i \in REF\_NO\_CONV \\ 1 & \text{If } i \in NEVER\_REF \end{cases}$$

where the groups *NEVER_REF*, *REF_CONV*, and *REF_NO_CONV* are defined in Table 5-3. This adjustment was done within sampling strata.

This approach has the effect of redistributing the sum of the weights of all the refusals not selected for conversion to the refusals that were selected for conversion, thereby setting the non-selected refusers sample weights to zero in the final combined file, but retaining the original sum of the weights in the remaining RDD sample. Note, too, that this step redistributes the weights to all sampled refusals regardless of final outcome, so that converted refusals, final refusals, and other outcomes will receive their appropriate share of the redistributed weights. The

records that were given a weight of 0 as a result of this adjustment were dropped from the weighting file and no further adjustments were made to them.

5.1.4 Multiple Telephone Adjustment

At the end of the screener interview, information about the existence of additional telephone numbers and their use in the household was collected. If the household had an additional telephone number that is used for residential purposes (not solely for business, fax or computer use, etc.), the household had a greater probability of selection because it could have been selected from the RDD frame through any of the additional telephone numbers in the household. If this is the case, the household weight was adjusted to reflect the increased probability of selection. The multiple telephone adjusted household weight, $HHA3W_i$, is computed as:

$$HHA3W_i = HHA3F_i * HHA2W_i ,$$

where $HHA3F_i$ is the multiple telephone adjustment factor computed as:

$$HHA3F_i = \begin{cases} 0.5 & \text{If the household has more than one residential telephone number} \\ 1 & \text{Otherwise} \end{cases} .$$

In this adjustment, it is standard practice to assume that there is at most one additional residential use telephone number in the household, since research has shown that there are not enough households having more than two such numbers to affect the weights if the exact number of phone lines is added to the formula.

5.1.5 Initial Person Weights

In the MATS 2007 RDD sample, one adult per household was selected for an interview. Adults were sampled with equal probability in households where only young adults (age 18-24 years) resided or only older adults (age 25 years and older) resided. In households where at least one young adult and one older adult resided ("mixed-age" households), the young adults were sampled at a rate four times that

of the older adults. The person level base weight reflects this differential probability of selection. The person baseweight, $PERIW_i$, is computed as

$$PERIW_i = \frac{1}{PROB_j} \cdot HHA3W_i,$$

where $PROB_j$ is the probability of selection of the *j*-th adult in household *i* computed as

$$PROB_j = \begin{cases} CNT_j^{-1} & \text{if } j \notin \text{mixed - age household} \\ p_O \cdot (OCNT_j)^{-1} & \text{if } j \in \text{mixed - age household and older adult selected} \\ p_Y \cdot (YCNT_j)^{-1} & \text{if } j \in \text{mixed - age household and young adult selected} \end{cases},$$

where CNT_j is the number of adults in the j^{th} household that is not mixed-age, p_O is the probability of selecting an older adult in a mixed age household, p_Y is the probability of selecting a younger adult in a mixed age household, and $OCNT_j$ and $YCNT_j$ are the number of older adults and number of young adults in the j^{th} mixed-age household, respectively. For MATS 2007, $p_O = 0.20$ and $p_Y = 0.80$.

5.1.6 Final Raked RDD Weight

The person level base weights were benchmarked to Minnesota adult population estimates, or control totals, defined by the 2005 American Community Survey. Benchmarking to control totals is a commonly used estimation procedure in which estimates are controlled to marginal population totals. The benchmarking method used on the MATS 2007 RDD weights was raking. Raking is an iterative procedure that can be thought of as multi-dimensional poststratification because the weights are poststratified to one set (a dimension) of control totals, then these adjusted weights are poststratified to another dimension. The procedure continues until all dimensions are adjusted. The process is then iterated until the control totals for all dimensions are simultaneously satisfied (at least within a specified tolerance). The raked weight, $RAKEDW_i$, can be expressed as

$$RAKEDW_i = PERIW_i \cdot \prod_{k=1}^{K} RAKEDF_{k_l},$$

where $RAKEDF_{k_l}$ is the raking factor for dimension *k*, level *l* which respondent *i* is in. For example, if the 4th dimension (*k* =4) is sex with two levels (*l*=1 for male and *l*=2 for female), then the raking factor for this dimension is $RAKEDF_{4_1}$ for the adult male. The raking factors are derived so the following relationship holds for every raking dimension *k*, and level *l*,

$$CNT_{k_l} = \sum_i \delta(k_l)_i \cdot RAKEDW_i ,$$

where CNT_{k_l} is the control total, and $\delta(k_l)_i = 1$ if the adult i is in level *l* of dimension *k* and zero otherwise. The MATS 2007 weights were raked to the five dimensions defined in Table 5-4. Raking to these five dimensions simultaneously controlled for gender, age, race, educational attainment, and location (inside or outside of the Twin Cities metropolitan area) RDD sample respondents were also asked to self-report whether they were members of Blue Cross. This information was collected for use in creating the composite weights for the combined RDD and Blue Cross samples, as the best, albeit imperfect, way of determining if someone in the RDD sample had a dual probability of selection because they were also a Blue Cross member. Using this self-report of Blue Cross membership, the RDD sample was also raked to a dimension defined by the membership counts within each age group on the Blue Cross frame.

5.1.7 Imputation of RDD Variables for Weighting

In order to poststratify the RDD weights, the missing values for a few variables were imputed.[17] The level of missingness for the variables to impute was extremely small, and resulted from the few responses of Don't Know or Refused to the needed variables; however, for post-stratification, every case must be assignable to the cells as defined by the raking dimensions. Due to the very small percentage of missingness, random allocation was used to impute the missing data. In this

[17] These imputed variables were used for weighting purposes only. The analysis will use the actual survey variables, with missing values retained. The names of the imputed variables are different from the actual survey variables.

Table 5-4. Description of variables used to define MATS 2007 RDD poststratification cells

Raking dimensions	Dimension definitions
Dimension 1	
Gender x Age groups	Male, 18 to 24 years
	Male, 25 to 29 years
	Male, 30 to 34 years
	Male, 35 to 44 years
	Male, 45 to 54 years
	Male, 55 to 64 years
	Male, 65 years and older
	Female, 18 to 24 years
	Female, 25 to 29 years
	Female, 30 to 34 years
	Female, 35 to 44 years
	Female, 45 to 54 years
	Female, 55 to 64 years
	Female, 65 years and older
Dimension 2	
Race/Ethnicity	Hispanic
	Non-Hispanic, White
	Non-Hispanic, African American
	Non-Hispanic, Asian
	Non-Hispanic, Native American, Pacific Islander
	Non-Hispanic, Other race, 2 or more races
Dimension 3	
Location x Collapsed Race	Inside Twin-Cities, Hispanic
	Inside Twin-Cities, Non-Hispanic, White
	Inside Twin-Cities, Non-Hispanic, African American
	Inside Twin-Cities, Non-Hispanic, Asian
	Inside Twin-Cities, Non-Hispanic, Native American, Pacific Islander, Other race, 2 or more races

Table 5-4. Description of variables used to define MATS 2007 RDD poststratification cells (continued)

Raking dimensions	Dimension definitions
	Outside Twin-Cities, Hispanic
	Outside Twin-Cities, Non-Hispanic, White
	Outside Twin-Cities, Non-Hispanic, African American, Asian, Native American, Pacific Islander, Other race, 2 or more races
Dimension 4	
Educational attainment x Age	Less than HS degree by the age groups in dimension 1
	HS degree or equivalent, by the age groups in dimension 1
	At least some college, by the age groups in dimension 1
	BS degree or above, by the age groups in dimension 1
Dimension 5	
Blue Cross frame count x Age	Self-reported Blue Cross member by the age groups in dimension 1
	Self-reported not a Blue Cross member

technique, a response is randomly selected from the observed distribution of the variable. So, for example, if gender is missing for a respondent and it is observed that 50 percent of respondents are male, then the missing value will have a 50 percent chance of being male. Table 5-5 lists the candidate variables for imputation as well as their rate of missingness. Age and gender did not have any missingness, while education, Hispanicity, race and location each had very low rates of missingness, ranging from 0.37 percent to 0.62 percent.

Table 5-5. Description of variables to be imputed for RDD raking

Imputed variable name	Description	Questionnaire item number	Percent missing
IMP_AGE	Imputed age	Derived from A5/A6	0.00
IMP_SEX	Imputed sex	J14	0.00
IMP_EDUC	Imputed highest level of education completed	Derived from J11	0.53
IMP_LOC	Imputed location	Derived from A7	0.62
IMP_RACE	Imputed race/ethnicity	Derived from J3, J4A, J4B, J4C, J4D, J4E, J4F	0.37

5.2 Blue Cross Sample

The RDD sampling is essentially a two-phase selection process. First, telephone numbers are selected. After a telephone number is called to determine if the number belongs to a residence, a household screener interview takes place where it is determined how many eligible adults are residing in the household. From those adults, a single respondent is selected to complete the interview. The Blue Cross sample process includes only one phase, since a sample of persons was selected from the Blue Cross membership list. This simplifies the weighting process compared to the RDD weighting process. The Blue Cross weights effectively started as person-level weights, whereas the RDD weights started as household-level weights and had to be converted into person-level weights after various adjustments. The initial Blue Cross weights are the inverse of the probability of selection for a sampled member as shown in Section 5.2.1. The initial Blue Cross weights underwent two subsequent adjustments; a deduplication adjustment similar to the RDD deduplication adjustment, and a raking adjustment that accounted for non-response and adjusted the weights so that they sum to the Blue Cross sampling frame counts. These two adjustments are discussed in Sections 5.2.2 and 5.2.3, respectively.

5.2.1 Initial Blue Cross Weights

The Blue Cross sample was drawn systematically from a frame comprised of Blue Cross members. Prior to selection of the sample, the frame was sorted by sampling stratum, age groups, gender, and location (inside or outside of Twin Cities metropolitan area). As described more fully in Section 2.2.2, the seven sampling strata were defined by plan type and age (Table 5-6). A higher rate of selection was used in the strata with young adults (Strata 5 - 7).

Table 5-6. Definition of MATS 2007 Blue Cross sampling strata

Stratum number	Blue Cross plan type	Age in years
1	PMAP	25 plus
2	MinnesotaCare	25 plus
3	Commercial plans	25 plus
4	Medicare Supplemental plans	25 plus
5	PMAP	18 – 24
6	MinnesotaCare	18 – 24
7	Commercial plans	18 – 24

Each sampled Blue Cross member in the Blue Cross sample was assigned an initial weight. The initial weight is computed as the inverse of the probability of selection of the member.

Let $BCIW_{hi}$ denote the initial weight for the *i*-th sampled Blue Cross member in the *h*-th stratum as defined in Table 5-6. Then the initial weight is given by

$$BCIW_{hi} = \frac{N_h}{n_h},$$

where N_h is the number of Blue Cross members in stratum *h*, and n_h is the number of members sampled in stratum *h*. As with the RDD discussion, the subscript h will be omitted in the description of the weights in the following sections as a notational convenience.

5.2.2 Sample Deduplication Adjustment

As fully described in Section 2.2.3, since the MATS 2007 sample design required sample from an RDD frame of telephone exchanges and a sample of persons from the Blue Cross frame, there was a possibility of a telephone number appearing in both the RDD sample and on the record of the Blue Cross sample member (duplicate number). Prior to the start of data collection, if a member was found to have a contact number (either provided on the Blue Cross administrative record or found by telephone look-up procedures) that matched a telephone number in the RDD sample, that member was dropped from the Blue Cross sample. The first adjustment to the Blue Cross base weights accounts for this subsampling of Blue Cross membership.

The Blue Cross duplicate-adjusted weight, $BCA1W_i$, is:

$$BCA1W_i = BC1F_i * BCIW_i$$

where $BC1F_i$ is the Blue Cross duplicate adjustment factor computed as:

$$BC1F_i = \begin{cases} \dfrac{\sum_{i \in DUP, NOT\_DUP} BCIW_i}{\sum_{i \in NOT\_DUP} BCIW_i} & i \in NOT\_DUP \\ 0 & i \in DUP \end{cases}$$

where the groups *DUP*, and *NOT_DUP* are defined in Table 5-7. This adjustment was performed within sampling stratum.

Table 5-7. Duplicate number groups

Duplicate group (DUPST)	Description
DUP	Blue Cross member contact number also in RDD sample
NOT_DUP	Blue Cross member contact number not in RDD sample

This approach has the effect of redistributing the sum of the base weights of all the *DUP* sample in the Blue Cross sample to the *NOT_DUP* sample in the Blue Cross sample, thereby setting the *DUP* sample weights to zero, but retaining the original sum of the weights in the remaining sample.

5.2.3 Final Blue Cross Weight

Similar to the RDD weighting process, the adjusted Blue Cross weights were raked to Blue Cross sampling frame totals. The Blue Cross raked weight, $BCRKDW_i$, can be expressed as

$$BCRKDW_i = BC2W_i \cdot \prod_{k=1}^{K} BCRKDF_{k_l} ,$$

where $BCRKDF_{k_l}$ is the raking factor for dimension *k*, level *l* which respondent *i* is in. As with the RDD raking factors, the Blue Cross raking factors are derived so the following relationship holds for every raking dimension *k*, and level *l*,

$$BCCNT_{k_l} = \sum_i \delta(k_l)_i \cdot BCRKDW_i ,$$

where $BCCNT_{k_l}$ is the control total, and $\delta(k_l)_i = 1$ if the Blue Cross member *i* is in level *l* of dimension *k* and zero otherwise. The MATS 2007 Blue Cross weights were raked to the two dimensions defined in Table 5-8. Raking to these two dimensions simultaneously controlled for gender, age, and location. Note that race and educational attainment cannot be used as they were for the RDD raking process because these are not available in the Blue Cross administrative record frame.

Table 5-8. Description of variables to be used when defining Blue Cross poststratification cells

Raking dimensions	Dimension definitions
Dimension 1	
Collapsed sampling stratum (plan type)	PMAP
	MinnesotaCare
	Commercial plans
	Medicare Supplemental plans
Dimension 2	
Gender x Location x Age	Male, In Twin Cities Area, 18-24 years
	Male, In Twin Cities Area, 25-29 years
	Male, In Twin Cities Area, 30-34 years
	Male, In Twin Cities Area, 35-44 years
	Male, In Twin Cities Area, 45-54 years
	Male, In Twin Cities Area, 55-64 years
	Male, In Twin Cities Area, 65+ years
	Female, In Twin Cities Area, 18-24 years
	Female, In Twin Cities Area, 65+ years
	Male, Outside Twin Cities Area, 45-54 years
	Male, Outside Twin Cities Area, 65+ years
	Female, Outside Twin Cities Area, 45-54 years
	Female, Outside Twin Cities Area, 65+ years

5.2.4 Imputation of Blue Cross Variables for Weighting

In order to poststratify the Blue Cross weights, the missing values for a few variables were imputed. Similar to the RDD rates of missingness, the level of missingness for the variables to impute for the Blue Cross sample was small and the reasons for missingness was the same. Due to the very small percentage of missingness, random allocation was used to impute the missing data. This technique is explained in section 5.1.7. Table 5-9 lists the candidate variables for imputation as well as their rate of missingness. While education and race/ethnicity could not be used for raking the Blue Cross sample to the Blue Cross frame control totals, it was necessary to impute them when missing on the Blue Cross sample because of the final step to combine the RDD and Blue Cross samples into a composite file that would be raked to the same set of control totals as were used for the RDD sample alone, which included totals for education and race/ethnicity categories. Age and gender did not have any missingness, while education, race/ethnicity and location each had very low rates of missingness, ranging from 0.38 percent to 0.50 percent.

Table 5-9. Description of variables to be imputed for MATS 2007 Blue Cross sample

Imputed variable name	Description	Extended interview items	Percent missing
IMP_AGE	Imputed age	Derived from A5/A6	0.00
IMP_SEX	Imputed sex	J14	0.00
IMP_EDUC	Imputed highest level of education completed	Derived from J11	0.50
IMP_LOC	Imputed location	Derived from A7	0.46
IMP_RACE	Imputed race/ethnicity	Derived from, J3, J4A, J4B, J4C, J4D, J4E, J4F	0.38

5.3 Composite (Blue Cross and RDD) Weights

To supplement the MATS 2007 RDD sample, the Blue Cross sample was combined with it into a single file resulting in 12,580 total completed interviews: 7,532 from the RDD component and 5,048 from the Blue Cross component. Since the Blue Cross frame count used for MATS 2007 sampling is 1.014 million and the April 2007 CPS population count of adults age 18 plus is about 3.8 million (a ratio of about 27 percent), a considerable portion of the RDD sample overlapped with the Blue Cross frame. These RDD cases are referred to as "overlap" sample. There is no way to determine with absolute certainty which respondents in the RDD sample comprise the overlap sample. In order to attempt to identify the overlap sample, the RDD respondents were asked if they have health insurance through Blue Cross and Blue Shield of Minnesota or Blue Plus (question I16 in the MATS 2007 extended interview). While imperfect, this is the most straightforward and reliable method, given the practical limitations of MATS 2007. There are three reasons why the chosen method is less than perfect:

1. a person could report having Blue Cross insurance when he or she does not, or could report not having Blue Cross insurance when he or she does; such respondent error could be due to faulty recall or false positives resulting from the pervasiveness of the Blue Cross brand as a near-generic term for health insurance;

2. timing differences, resulting from defining Blue Cross membership for MATS 2007 purposes as the membership list as of January 1, 2007, while the data collection occurred in February – June, 2007; and

3. definitional differences, resulting from excluding certain plan types and employees from the Blue Cross frame for MATS 2007 purposes (see Section 2.2.1), while excluded members sampled through the RDD process could validly report themselves as Blue Cross members in response to RDD question I16, since it was a simple question about general Blue Cross membership, without qualification.

There are in theory other potential means of determining the overlap sample, for example, identifying the phone numbers of the RDD respondents that are also on the Blue Cross administrative record frame. Once a phone number has been

matched, demographic characteristics such as gender and age as reported in the RDD interview could be compared to the membership data on the frame with the same phone number. This can establish the "likelihood" that the person is actually on the frame or not. This approach was avoided because of the observed incompleteness of phone numbers on the Blue Cross frame and the observed inaccuracy of those numbers experienced through phone number look-ups and actual data collection outcomes. A further complication of this approach is that, due to family membership, more than one person may have the same phone number on the Blue Cross frame. More certain methods that used to be available to survey researchers – such as asking for the RDD respondent's Blue Cross membership number and then checking it against the Blue Cross member list – would have been regarded with suspicion by the respondents, would have yielded a low cooperation rate, would not have been permitted by Blue Cross, and would have created difficulty securing approval by a required Institutional Review Board in light of increasing concerns about individual privacy, health information confidentiality, and identity theft.

For these reasons, only question I16 was used to identify the overlap sample. As described below, certain metrics were applied to determine the face validity of this approach.

Once the overlap sample was identified, a composite weight was created for the identified RDD overlap and Blue Cross sample cases using the principles of composite estimation so that the combined file can be used for analysis.

In composite estimation, the estimates being combined are assumed to be independent, unbiased estimates of the same population parameter. In other words, the Blue Cross sample and the RDD overlap sample cases that meet the Blue Cross sample eligibility criteria theoretically should be representative of the same population. Inspecting the overlap sample and the Blue Cross sample can determine if this assumption was met, by comparing (weighted) estimates of variables from both samples, such as Blue Cross membership counts, and distributions of age, gender, self-reported race, location and prevalence of current and former smokers. A special RDD weight was created for this comparison that removed race and education from the poststratification procedure. This was done in

order to make the estimates more comparable, since these demographics do not exist in Blue Cross records and therefore could not be controlled for when post-stratifying the Blue Cross sample. Table 5-10 shows the comparisons of the distributions of these variables for the Blue Cross sample and for the overlap portion of the RDD sample that indicated Blue Cross membership, referred to as the RDD-Blue Cross sample.

Table 5-10. Weighted distributions of demographic variables common to the Blue Cross List and RDD-Blue Cross samples

Variable	Blue Cross List Sample %	RDD-Blue Cross Sample %	Difference %
Total	100.0	100.0	
Smoking Status			
Current	10.8	12.4	1.6
Former	27.9	27.6	-0.2
Never	61.3	60.0	-1.4
Age			
18-24 years	11.5	11.3	-0.1
25-44 years	31.9	31.4	-0.5
45-64 years	35.6	36.7	1.1
65 plus years	21.0	20.6	-0.4
Sex			
Male	46.7	48.8	2.1
Female	53.3	51.2	-2.1
Location			
In Twin Cities	48.4	47.8	-0.7
Outside Twin Cities	51.6	52.2	0.7
Race			
Hispanic	1.6	3.0	1.4
White, Non-Hispanic	94.0	92.2	-1.9
Black, Non-Hispanic	0.7	1.9	1.2
Other	3.7	2.9	-0.8
Education			
Less than HS	5.0	5.8	0.8
HS or equivalent	26.6	27.0	0.5
Some college	31.5	31.4	0.0
BS degree or more	37.0	35.7	1.2

A related assumption needed for the creation of the composite weights is that all overlap sample cases can be identified. It is, therefore, necessary to make some allowance for the sources of error or differences resulting from the self-report method used to identify those cases among the RDD respondents, as described above.

5.3.1 Calculation of MATS 2007 Composite Weights

The composite weight, w_{comp}, for each respondent in the Blue Cross sample and the overlap sample was calculated as

$$w_{comp} = \bar{\lambda}_j w_1 + (1 - \bar{\lambda}_j) w_2 .$$

where

$\bar{\lambda}_j =$ the lambda value for the j^{th} age group, where j = 1, 2, 3, 4.
$w_1 =$ original List sample weight, and
$w_2 =$ original overlap sample weight.

As seen below, the lambda values are a ratio of variances of survey estimates. Since the variances for the survey estimates were variable by age group (18-24 years, 25-44 years, 45-64 years, and 65 years and older), the calculations of the composite weights were done within these age groups. However, as a notational convenience, the rest of the calculations appear without the subscript indicating the j^{th} age group. The weighted average of lambda was calculated as

$$\bar{\lambda} = \frac{\lambda_1 n_1^{eff} + \lambda_2 n_2^{eff}}{\sum n_i^{eff}}, \text{ where}$$

$$\lambda_1 = \frac{\sigma_{12}^2}{\sigma_{11}^2 + \sigma_{12}^2},$$

$\sigma_{11}^2 =$ variance of the smoking rate estimate from the Blue Cross sample, and
$\sigma_{12}^2 =$ variance of the smoking rate estimate from the overlap sample.

$$\lambda_2 = \frac{\sigma_{22}^2}{\sigma_{21}^2 + \sigma_{22}^2},$$

$\sigma_{21}^2 =$ variance of the former-smoking rate estimate from the Blue Cross sample, and
$\sigma_{22}^2 =$ variance of the complement of the former-smoking rate estimate from the overlap sample.
n_1^{eff} and n_2^{eff} are the effective sample sizes of the Blue Cross and overlap samples, respectively.

Therefore, $\bar{\lambda}$ is a weighted average of the lambda values, and by using the weighted average, the composite weight gives increased weight to the estimate with the smaller variance. Estimates of proportions and their variance were computed for smoking status for the Blue Cross sample and the overlap portions separately.

To create the composite weight, the sample weight from the Blue Cross sample was multiplied by $\bar{\lambda}$ and for the overlap cases, the sample weight was multiplied by a factor of $1-\bar{\lambda}$. The remaining RDD sample cases were assigned a composite weight equal to their original RDD sample weight (i.e., their original weights were multiplied by a factor of unity).

After applying the lambda factors to the weights to create composite weights, a final repetition of the raking process was performed to bring the final composite weights into agreement with the ACS control totals used for the original RDD weights. The same dimensions used during the RDD raking process were used for the composite raking.

5.3.1.1 A Note of Caution in Regard to the Weighted RDD and Weighted Composite MATS 2007 Samples

It is important to note that, while the weighted MATS 2007 RDD sample and the weighted MATS 2007 Composite sample are both random samples of the adult Minnesota population and can produce unbiased estimates, each is a sample with associated sampling error. The point estimate of a parameter calculated using the weighted RDD sample will not necessarily be identical to the point estimate of the same parameter using the Composite sample. It is the intention and wish of the survey sponsors and designers that only the Composite sample be used for all MATS 2007 analyses and reports that pertain to the Minnesota population.

5.4 Reweighting of MATS 1999 and MATS 2003

It is well established that smoking behavior and other health-related behavior and risk activities are correlated with certain common demographic characteristics, including gender, age, and education. As described in Section 5.1.6, the MATS 2007 weighting plan incorporated these and several other variables into the raking

design for the final weights. Exploration of the MATS 2007 data to develop the final weighting classes confirmed that smoking status was sensitive to educational level (the higher the educational level, the lower the smoking prevalence). Because MATS 2007 – and most RDD surveys in general – over-represented the more educated and under-represented the less educated, excluding education as an adjustment factor would tend to understate the smoking prevalence rate for the full population.

While MATS 2007 utilized educational status in creating the weights, MATS 1999 and MATS 2003 did not. Investigators from the University of Minnesota had weighted MATS 1999 and MATS 2003 in accordance with generally accepted practices, such as CDC's BRFSS and other statewide tobacco surveys. These surveys generally post-stratify only on age and gender. However, with the incorporation of education into the weights of MATS 2007, the absence of education in the weights of MATS 1999 and MATS 2003 presented a very real implication for the comparability of the estimates across the three MATS survey rounds for purposes of trend analysis, one of the major analytic objectives of the MATS series. Westat engaged in considerable exploration of the effects of incorporating educational status into the existing MATS 1999 and 2003 weights, then prepared a detailed memo laying out the findings of this exploration and possible approaches to reweighting the MATS 1999 and 2003 data to make them more comparable to MATS 2007 for the trend analysis. The MATS Advisory Panel reviewed the memo and considered the options, deciding to proceed with the MATS 1999 and 2003 reweighting for use in the trend analysis. These weights would be used for the present trend analyses covering MATS 1999 through MATS 2007 and would be used in trend analyses in future MATS surveys. Therefore, estimates from MATS 1999 and 2003 using these adjusted weights may vary slightly from estimates reported in previous years.

5.4.1 Reweighting Method

The reweighting of MATS 1999 and MATS 2003 can be thought of as a supplemental calibration of the original weights that were calibrated when each of the survey was completed. Rather than recreating the respective baseweights and adjusting those to create a new final analysis weight, the reweighting used the

existing final, calibrated weights as input weights, then recalibrated these weights using raking dimensions similar to those used in MATS 2007.

To recalibrate the MATS 1999 and MATS 2003 weights, Westat developed control total dimensions that mirrored as closely as possible the control total dimensions used in MATS 2007, in an effort to create poststratification dimensions defined the same way as they were defined in MATS 2007 (see Table 5-4 in Section 5.1.6).

As with MATS 2007, the MATS 2003 reweighting used ACS data, specifically 2003 ACS population totals. However, since ACS data was not available for the MATS 1999 data collection year, Census 2000 Long Form data was the closest match.

In addition, there were a couple of required modifications to the poststratification dimensions due to differences in MATS 1999 and 2003 relative to MATS 2007. Since RDD sample interviews did not collect information on Blue Cross membership in MATS 1999 and 2003, the "Blue Cross Frame count" dimension was not included in the reweighting. Secondly, the detailed definitions within the other dimensions were adjusted through collapsing to reflect the smaller sample sizes from MATS 1999 and 2003 relative to the MATS 2007 sample size.

5.4.2 Reweighting Results

Overall, the outcomes were as expected. For MATS 1999, the estimated smoking prevalence rate increased from 20.3% as previously reported to 22.1%. The MATS 2003 rate increased from 18.0% as previously reported to 19.1%.

Table 5-11 presents the results of the MATS 1999 and MATS 2003 reweighting for the Education variable. For comparison purposes, the original weighting results for MATS 1999 and 2003 are provided along with the MATS 2007 weighting results.

Table 5-11 shows that the increase in the estimated smoking rates for MATS 1999 and 2003 are mainly driven by the reduced proportion of the weighted sample coming from the "BS or more" (college degree) subgroup, which has a much lower smoking rate than those with lower levels of educational attainment. For example, from MATS 1999, the "BS or more" smoking rate of 10.4% is associated with only 25.1% of the population (Reweight) instead of 35.5% (Original) of the population.

Again, the overcoverage of the population with higher levels of education was artificially driving the smoking rates down with the original weighting. The reweighting has corrected for this and further reduced the bias.

The overall trend for adult smoking prevalence after reweighting MATS 1999 and 2003 is 22.1% (MATS 1999), 19.1% (MATS 2003), and 17.0% (MATS 2007).

Table 5-11. MATS 1999 and 2003 population distributions and smoking status rates before and after reweighting by education

		MATS I				MATS II				MATS III	
		Original		Reweight		Original		Reweight			
EDUCATION	Smoking Status	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
HS grad or less	Current	333,993	27.0%	405,664	27.0%	313,005	25.7%	364,124	24.8%	348,413	24.6%
	Former, Never	905,137	73.0%	1,098,362	73.0%	904,913	74.3%	1,105,305	75.2%	1,067,283	75.4%
Total		1,239,130	36.2%	1,504,026	41.6%	1,217,917	32.9%	1,469,429	40.1%	1,415,696	38.0%
Some College	Current	234,822	24.2%	298,183	24.8%	229,219	20.1%	241,310	20.6%	221,639	17.8%
	Former, Never	735,526	75.8%	906,582	75.2%	913,079	79.9%	931,306	79.4%	1,023,784	82.2%
Total		970,347	28.3%	1,204,765	33.3%	1,142,298	30.9%	1,172,616	32.0%	1,245,423	33.4%
BS or more	Current	127,547	10.5%	94,311	10.4%	124,148	9.3%	96,005	9.4%	63,521	6.0%
	Former, Never	1,089,288	89.5%	813,007	89.6%	1,213,738	90.7%	928,736	90.6%	1,000,410	94.0%
Total		1,216,835	35.5%	907,319	25.1%	1,337,886	36.2%	1,024,741	27.9%	1,063,931	28.6%
Overall	Current	696,361	20.3%	798,158	22.1%	666,372	18.0%	701,439	19.1%	633,573	17.0%
	Former, Never	2,729,951	79.7%	2,817,951	77.9%	3,031,729	82.0%	2,965,347	80.9%	3,091,477	83.0%
Total		3,426,312	100.0%	3,616,109	100.0%	3,698,101	100.0%	3,666,786	100.0%	3,725,051	100.0%

6. Comparability with Previous MATS Surveys and Limitations of the Data

It is helpful for users of the MATS 2007 data to be aware of the survey methods described in this methodology report, when analyzing the data, interpreting the findings, reading reports, and applying the results to historical assessment and planning for the future. It is likewise helpful for them to be informed of any known or potential limitations that apply to the use of the data. Finally, when comparing results of MATS 2007 to those of previous MATS surveys, it is important to consider methodological factors that may affect the comparability of the data from one round to the next. This report focuses mainly on comparability of MATS 2007 with the immediately previous round, MATS 2003, but also examines comparability across all three rounds when relevant and feasible.

6.1 Comparability Issues

Several issues related to comparability, or factors affecting it, have been discussed in the relevant sections of this report. Briefly, these are aspects of the following issues:

Sampling issues:

- Oversampling African Americans
- Oversampling young adults in the RDD sample
- Blue Cross sampling strata and selection probability
- Oversampling young adults in the Blue Cross sample
- Criteria for fielding Blue Cross sample cases and defining sample denominator
- Handling the deduplication of RDD sample phone numbers and phone numbers of sampled Blue Cross members

Weighting issues:

- Weighting design
- Combining and weighting the RDD-Blue Cross sample

Data collection issues:

- Screener refusal conversion
- Home-based interviewers

Questionnaire design and specification issues:

- Questionnaire content
- Questionnaire skip patterns

6.1.1 Sampling

6.1.1.1 Oversampling African Americans in RDD Sample

MATS 2007 RDD oversampled African Americans for the first time. This not only yielded a higher proportion of African Americans in the sample, but it also yielded a higher proportion of non-African Americans who live in geographical areas with high concentrations of African Americans. While the weighting and raking adjust for much of the difference resulting from oversampled populations and the natural distribution of all Minnesotans on major demographic characteristics, they do not necessarily adjust fully for characteristics that are not associated with those characteristics. To the extent that the higher proportion from certain geographical areas may produce overall survey estimates that are different from the estimates that would have been obtained without oversampling, the sample design would affect comparability with previous rounds of MATS, but the expected effect would be minimal.

6.1.1.2 Oversampling Young Adults in RDD Sample

MATS 2003 and MATS 2007 both oversampled young adults in the RDD survey. MATS 1999 did not over sample young adults. MATS 2003 oversampled them at a rate of nine times the other adults and MATS 2007 at a rate of four times the others.

The lowering of the MATS 2007 oversampling rate should very slightly reduce sampling variance in the RDD and combined files and marginally improve the precision of overall estimates, but would not affect the point estimates themselves. The actual proportion of young adults in the respective samples could affect overall point estimates, to the extent that differences between young and older adults on the characteristic in question are not adjusted back to their natural distribution by the weighting process.

6.1.1.3 Blue Cross Sampling Strata and Selection Probability

MATS 2007 and MATS 2003 both oversampled young adults in the Blue Cross sample. MATS 2003 also oversampled the other adult members of certain Blue Cross health plan types, while MATS 2007 selected all other adult members with equal probability across plan types. The equal selection probability for most of the MATS 2007 Blue Cross sample is likely to have somewhat reduced the sampling variance and therefore improved the precision of the resulting survey estimates.

6.1.1.4 Oversampling Young Adults in Blue Cross Sample

MATS 2003 and MATS 2007 both oversampled young adults in the Blue Cross survey. MATS 2003 oversampled them at varying rates by plan stratum, while MATS 2007 sampled them at a lower, uniform rate across strata. It is difficult to assess the affect on precision of lowering and standardizing the MATS 2007 oversampling rate, since the sampling rates of the older adults varied quite a bit in MATS 2003, and some were at a higher rate and some at a lower rate than the young adults. As with the RDD oversample of young adults, the actual proportion of young adults in the respective samples could affect overall point estimates in the Blue Cross and combined file data, to the extent that differences between young and older adults on the characteristic in question are not adjusted back their natural distribution by the weighting process.

6.1.1.5 Criteria for Fielding Blue Cross Cases and Defining Sample Denominator

The MATS 2007 Blue Cross sample included every eligible case, even those that had no phone number, and treated the latter as non-response in calculating response rates and in the weighting and raking process. MATS 2003 excluded such cases,

and also excluded Blue Cross cases with phone numbers in Blue Cross records that were not confirmed as having a listed number through a phone number database look-up process. As a result, the MATS 2003 responding cases may have been easier to locate. Willingness to have a listed phone number is also associated with a propensity to respond to telephone surveys specifically. MATS 2007 also included every eligible sampled case in the denominator for calculating the Blue Cross response rate, while MATS 2003 excluded a large number of cases dropped from the fielded sample. These differences might have tended to lower the operational yields of the MATS 2007 survey compared to MATS 2003, and also to have reduced the potential bias associated with sample design and fielding protocol (MATS 2007 is completely unbiased in this regard).

6.1.1.6 Deduplication of RDD Sample Phone Numbers and Phone Numbers of Sampled Blue Cross Members

MATS 2007 dropped Blue Cross sample members whose phone number was also found in the RDD sample; MATS 2003 did the opposite. As discussed in Section 2.2.3.2, the MATS 2007 approach reduces potential bias.

6.1.2 Weighting

6.1.2.1 Weighting Design

MATS 2007 used a similar weighting approach to MATS 2003, but the specifics of the design differed somewhat. The main differences were in the post-stratification/raking design.

MATS 2003 post-stratified the RDD sample to two age groups (18-24-year-olds and 25 and older), gender, and two geographical groups. MATS 2007 raked the RDD sample to age, gender, geography, race, and education. MATS 2007 used more levels of age and crossed certain of the variables in the raking process. The MATS 2007 raking approach allowed finer degrees of adjusting the responding sample to the overall Minnesota population. Most importantly, the addition of education as a raking variable had a major impact on smoking prevalence estimates and likely on other smoking estimates, since it is associated with smoking behavior. The absence of this variable in the MATS 1999 and 2003 weighting process made trend comparisons across the three surveys less consistent. For this reason, the MATS

2007 Advisory Panel decided to reweight the MATS 1999 and 2003 samples by adding education as a raking dimension, to make the survey estimates more comparable, as described in Section 5.4.

MATS 2003 post-stratified the Blue Cross sample on the seven sampling strata, gender, and three geographical levels. MATS 2007 raked to four plan type sampling strata and to multiple levels of combinations of age, gender, and geography. As in the RDD, the MATS 2007 Blue Cross raking approach allowed finer degrees of adjusting the responding sample to the overall Blue Cross population. The addition of age as a raking variable may have improved estimates of smoking prevalence and other smoking estimates, since it is associated with smoking behavior.

6.1.2.2 Combining and Weighting the RDD and Blue Cross Samples

MATS 2007 combined and weighted the RDD and Blue Cross samples in a different manner from MATS 2003.

MATS 2003 and MATS 2007 used different approaches to combine the RDD and Blue Cross samples, to account for the dual probability of selection among Blue Cross members and to create the composite weights for the combined file. MATS 2003 accounted for the dual probability by matching the individual Blue Cross respondent sample cases to RDD sample cases, using various demographic and other characteristics such as type of health insurance and possession of a listed telephone number of the respondents. More than one Blue Cross case could match an RDD case. When the matching step was done, the RDD sample weight was divided between the matched Blue Cross respondent and the RDD respondent in proportion to their initial weights. MATS 2007 used a new question added to the MATS 2007 questionnaire to determine the RDD respondent's self-reported membership in a Blue Cross or Blue Plus health plan and assign those RDD cases who reported membership to an "overlap" sample group. MATS 2007 then calculated a factor to apportion the overlap RDD cases' weights between the overlap cases and the Blue Cross cases. The apportioning factor took the variance of smoking rate estimates into account, and it was calculated and applied

independently for each of four age groups, since smoking behavior is associated with age.

MATS 2003 did not merge the Blue Cross PMAP stratum cases into the combined file. As stated directly above, one of the criteria used to match individual Blue Cross sample cases to RDD sample cases in order to apportion weights was type of health insurance coverage. However, there were concerns regarding significant population differences between the Blue Cross PMAP population and the statewide Medical Assistance population. For example, the majority of statewide disabled Medical Assistance participants are not enrolled in PMAP. This type of systematic selection into Blue Cross PMAP based on variable(s) not explicitly used to match the cases would have made the matching process invalid. Because MATS 2007 did not combine the samples using this RDD-Blue Cross case-matching process, MATS 2007 was able to include all Blue Cross cases in the final merged sample.

Finally, after combining the two files, MATS 2007 applied the same raking process to the combined file as to the RDD file, to further fine tune the weights along demographic dimensions, especially those such as age and education that are associated with smoking behavior.

MATS 1999 was exclusively an RDD sample.

The effects of these various differences in the weighting process and file composition of the final files used for calculating the survey estimates in the three MATS surveys on comparability are not clearly determinable. MATS 2007 should have more sensitivity to differences in smoking-related estimates overall and among demographic subgroups, and observed differences across time could be somewhat attributable to the weighting process. The reweighting of MATS 1999 and 2003 should ameliorate such artificial differences to a large extent.

6.1.3 Data Collection

All of the MATS survey data were collected using computer-assisted telephone interviewing, and adhered to the CDC BRFSS data collection protocol. They all used standard survey research interviewer training and interviewing protocols.

6.1.3.1 Screener Refusal Conversion

MATS 2007 adopted refusal conversion of randomly subsampled screener refusals and accounted for them in the weighting process. MATS 1999 and 2003 did not subsample refusals. The additional subsampling should produce a very slight increase in the sampling variance and therefore in the statistical precision of the estimates, but not in the point estimates themselves. Like the previous MATS surveys, MATS 2007 attempted to convert 100% of those who initially refused to respond to the extended interview. None of the MATS surveys attempted to convert refusals that were classified as "adamant" or "hostile."

MATS 1999 did not use refusal conversion letters, MATS 2003 sent them at the end of data collection to a subset of the remaining screener refusals, and MATS 2007 sent them to all screener refusals that had a matched address associated with the phone number.

6.1.3.2 Home-based Interviewers

MATS 2007 employed a mix of home-based interviewers and center-based interviewers. MATS 1999 and 2003 employed only center-based interviewers. There is no basis to ascribe any incomparability to the survey estimates resulting from the data collected under these two staffing models.

6.1.4 Questionnaire Design and Specification

There are two main areas where questionnaire design may affect comparability. The first area is the questionnaire content, which refers to the selection of questions, response categories, and the formulation of their specific wording and ordering. This area also includes the introductory text and transition phrases, as well as prompts, probes, and instructions to be used by the interviewers.

The second area is the determination of which respondents are administered each question and, for some questions, an alternative, more suitable phrasing of the question. This concept is commonly referred to as the "skip patterns" for the questionnaire. Some questions will not apply to certain groups because of who they are (questions about quitting smoking are not relevant to never smokers) or how they answered a specific question (if a person has not experienced any anti-

smoking ads, it is logical to skip the following question about the frequency of such experiences). The group who receives each question is usually referred to as the "base" for that question. In administering the questionnaire, the skip instructions determine and control who receives each question. All of these concepts are captured in detailed questionnaire specifications and in the programming instructions for CATI questionnaires.

6.1.4.1 Questionnaire Content

Section 3.1 described the general questionnaire design process and general issues and factors considered in formulating the question items, wording, and response categories. As noted there, and as elaborated in the Minnesota Adult Tobacco Survey 2007 Comparability Report and its item-by-item crosswalk comparison between MATS 2003 and 2007, there are a number of questions that appear in only one of the two most recent MATS questionnaires. Such questions, by definition, have no comparable items for trend comparison across the survey rounds. When previous items were omitted from the MATS 2007 questionnaire, the resulting absence of trend data was consciously anticipated in the design, either because the items were no longer of interest or had not been useful in the past, or because some items needed to be eliminated as a trade-off to accommodate new items. In addition, when new items were added, it was because of new or changing research objectives. While historical trends cannot currently be analyzed for new MATS 2007 items, MATS may choose to retain such items in the future and monitor the trend from MATS 2007 forward.

In addition to noting questions that exist only in MATS 2003 or MATS 2007, the MATS 2007 comparability report and questionnaire crosswalk fully document the various wording changes to clarify meaning, add or improve response categories, or simplify administration of questions appearing in both rounds. For the most part, the changes were minor and would not hinder meaningful comparisons across time. Appendix F discusses the nature and possible effect on comparability of significant MATS 2003 questionnaire items that were modified for MATS 2007. The question numbers refer to the MATS 2007 questionnaire attached as Appendix A.

6.1.4.2 Skip Patterns

6.1.4.2.1 Skip Pattern in MATS 2007 Compared to MATS 2003

A general change throughout the MATS 2007 questionnaire is that most questions that were in some way limited to 18- to 24-year-olds in MATS 2003 were not limited by age in MATS 2007. For example, a number of smoking questions were skipped in MATS 2003 for non-established smokers (smoked less than 100 cigarettes) over the age of 24, but asked of all non-established smokers in 2007, regardless of age. The implication of this is that trend comparisons using such questions have to be limited to the base defined by the subset of respondents common to both MATS 2003 and MATS 2007.

Appendix G discusses the nature and possible effect on comparability of significant MATS 2003 skip patterns that were changed for MATS 2007. The question numbers refer to the MATS 2007 questionnaire attached as Appendix A.

Especially significant were two instances where the process of defining MATS 2007 skip patterns and specifying them for the CATI questionnaire inadvertently excluded respondents from a series of questions that were needed for several desired analyses of both the MATS 2007 cross-sectional data and for trend comparisons with previous MATS.

Stages of Change. The first of these oversights impacted the ability to report on the standard concept of stages of change, a construct that classifies current smokers as to their readiness to quit smoking and former smokers as to the durability of their currently quit status. Questions E16 and E17 are needed to classify the stage of change for current smokers. The intention, consistent with MATS 2003, was for all current smokers, and all former smokers who smoked in the past 30 days, to be administered these questions, and for all others to skip to question G1. However, a skip instruction at an earlier point in the questionnaire (following E10) caused current smokers who had not made a quit attempt in the past 12 months to skip from that point to G1, thereby excluding them from the base for the stages of change questions. This affected approximately 46% (unweighted) of the smokers in the sample (48% weighted). As a consequence, MATS 2007 is unable to report on the stage of change for all current smokers in 2007 or to use this variable as an

independent variable for other outcomes where all current smokers are the population of interest. Likewise, it is not possible to extend the trend from MATS 2003 for any analysis that requires stage of change for all current smokers.

This problem did not affect determining the stage of change for former smokers or for the large subset of current smokers who had made a quit attempt in the past 12 months.

Self-efficacy for quitting and beliefs about stop-smoking medications. The same skip instruction that adversely affected collection of the stage of change data also affected the two questions about self-efficacy for quitting (E18) and beliefs about stop-smoking medications (E19a –e). These data are unavailable in MATS 2007 for current smokers who did not make a quit attempt in the past 12 months. As a consequence, MATS 2007 is unable to report on them for all current smokers in 2007 or to use any of these variables as an independent variable for other outcomes where all current smokers are the population of interest. Likewise, it is not possible to extend the trend from MATS 2003 for the self-efficacy variable. The questions about beliefs about stop-smoking medications are new to MATS 2007.

Like stage of change, self-efficacy is most useful when looked at across all smokers, not only those with a quit attempt in the past 12 months. As a result, the MATS 2007 report, *Creating a Healthier Minnesota: Progress in Reducing Tobacco Use,* does not examine either of these constructs.

In comparison, analyzing their various beliefs about stop-smoking medications is useful even when limited to current smokers with a quit attempt in the past 12 months. Perceptions of quitting assistance may affect smokers' interest in them or willingness to use them. Further, successful quitters have usually made multiple quit attempts before being successful. Those current smokers who had recently tried to quit are the most likely to make another attempt soon, and supporting this group of smokers in future attempts to quit is critical to Minnesota's tobacco control efforts.

6.1.4.2.2 Skip Pattern Issue Common to Both MATS 2003 and MATS 2007

Use of stop-smoking medications and other forms of stop-smoking assistance (E4a – E4f, E11 – E14) for young adult experimenters ("unrecognized smokers"). MATS 2003 and 2007 have put considerable focus on young adults aged 18 to 24. MATS investigators have developed the concept of "unrecognized smokers" in this age group. This set of young adults has not met the traditional threshold of having smoked 100 cigarettes that is used to define established smokers, but they show current smoking activity by having smoked in the last 30 days. This behavior may be indicative of experimentation that may lead to becoming an established smoker, so MATS analyzes young adult smokers by combining the established smokers and the "unrecognized smokers." MATS 2003 and MATS 2007 collected data about use of the various forms of stop-smoking assistance from all established smokers. However, both MATS 2003 and MATS 2007 skipped these questions for the unrecognized smokers. As a consequence, neither MATS 2003 nor MATS 2007 was able to report on these items for all young adult smokers as defined for MATS. MATS 2007 adopted the approach of examining them for young adults who met the traditional definition of established smoker, which was not consistent with most analyses that focused on young adult smokers.

6.1.4.3 Restructuring of Questions about Use of Stop-Smoking Medications and Behavioral Counseling

Two important areas covered by MATS are smokers' use of various types of stop-smoking medications and various forms of behavioral counseling to assist them in their attempts to quit smoking. MATS 2007 and MATS 2003 asked the same series of questions about each of these two topics, but MATS 2007 changed the way in which they were structured. The MATS 2007 questions are E4a-E4f (use of specific stop-smoking medications during most recent quit attempt) and E11-E15 (use of specific forms of behavioral counseling). Each of these questions was asked of everyone with an applicable quit attempt in MATS 2007. In MATS 2003, a gateway question first asked generally if the quitter had used any medications, then asked about specific medications only of those who responded positively. The same approach was then used for the behavioral counseling questions.

These modifications are essentially structural changes, eliminating the two MATS 2003 gateway questions and the skipping of the specific follow-on questions when the response to the gateway was negative. However, this structural change could have resulted in some additional positive, valid responses to the specific medications when they were presented individually, since respondents may not have been aware of all possibilities when answering the gateway question negatively in MATS 2003 or would not have thought of the medication until prompted about it by the specific question. Alternately, a MATS 2003 respondent could have self-defined medications in ways other than the commonly accepted medications intended by the survey designers. If so, this could have slightly inflated the MATS 2003 estimates of those who used any medication, compared to MATS 2007. The same considerations apply to the behavioral counseling questions.

There was the possibility that some small amount of any differences observed in the estimates produced from MATS 2007 compared to MATS 2003 could be due to this revised structure. Nonetheless, the improved sensitivity of asking each item separately was designed to improve the accuracy of the data about the use of medications and behavioral therapy for 2007 and in the future.

6.2 Potential Limitations of the Data

All of the Minnesota Adult Tobacco Surveys yield data that provides highly accurate and detailed representations of the smoking-related attitudes, beliefs and behaviors of Minnesota's adult residents at various points in time. However, statistics produced from sample surveys are subject to two general types of error, technically referred to as sampling error and nonsampling error. The term "error" does not refer to a mistake or a known error but to the fact there may be some difference between the survey statistic and the actual statistic for the entire population that the sample survey is meant to represent. It is for this reason that statistics produced from a sample are referred to as "estimates": they estimate what the actual statistics are for the entire population, or for any subgroup in the population.

6.2.1 Sampling Error

Sampling error is a purely statistical phenomenon, resulting from the fact that the data are collected from a sample that represents the entire population, rather than from everyone in the population, as in the case of a census. Sampling "error" is a technical term; it does not refer to any known error, but rather to the fact that an estimate produced from a sample has some amount of uncertainty associated with it.

It is possible to quantify the uncertainty of an estimate produced from the survey sample data, to the extent that the uncertainty is caused by the use of sample with a known probability of being selected for the survey. There is no one number that can be assigned to every survey statistic to indicate the uncertainty; rather, it depends on the type of statistic (percentage, mean, ratio, difference, etc.), the size of the sample used to calculate the estimate, and the effects of complex sampling designs such as those used for MATS.

Common measures of uncertainty include standard errors and confidence intervals. The MATS technical reports utilize confidence intervals, which express the likely range of the actual value of a population statistic, around the "point" estimate produced from the survey data. For example, the statement that MATS 2007 found the 2007 smoking prevalence among adult Minnesotans to be 17.0±1.4 percent means the expectation is that the true value falls somewhere within the confidence interval ranging from 15.6 percent to 18.4 percent. The confidence interval is commonly expressed as a "half-width," plus or minus around the point estimate, as in this smoking prevalence example. Like nearly all sample surveys, MATS reports the 95% confidence interval, which means that there is a 95% certainty that the interval for any given estimate contains the true value.

All statistics presented in the MATS technical reports utilize weighted data. The survey weights reflect the complex MATS 2007 sample design, as described in Chapter 5. This means that the reported statistics are reflective of the entire population or subgroup for which they are calculated. The weighted estimates for the MATS 2007 technical report and their associated confidence intervals were all

calculated using SAS and SUDAAN, widely used statistical software packages that account for the complex sample design and sample weights.

6.2.2 Potential Sources of Nonsampling Error

As in the case of sampling error, it is nearly impossible for a survey to avoid other sources of error. Unlike sampling error, it is not typically possible to quantify potential nonsampling errors in a specific survey.

6.2.2.1 Frame Coverage Issues

In addition to the sampling error that is common to all sample surveys, MATS 2007 was also subject to a form of nonsampling error known as coverage error. All survey samples use a "frame" from which to draw the sample. Ideally, the frame "covers" the same population about which the survey seeks to provide information, but frames seldom perfectly cover the population. Those in the population who are not covered by the frame may be different from those who are covered by it, in terms of the characteristics, behaviors, attitudes, and beliefs that the survey addresses. The greater this difference (if any), the greater the likelihood that there is some error in the reported statistics, in terms of their accurately reflecting the entire population.

In the case of RDD surveys, which use landline telephone numbers as the frame, the coverage historically was above 95 percent, since less than 5 percent of the population was without telephone service at any point in time. In recent years, many households have begun using their cell phones as their household telephone line, abandoning totally their regular landline telephones. As of the first half of 2007, it was estimated that, nationally, 12.6 percent of adults lived in households that used cell phones exclusively. Such cell phone-only households are not covered by RDD surveys. Further, the prevalence of cell phone-only households is considerably higher among younger adults (18-29), African Americans, and Hispanics.[18] Further still, members of such subgroups in households that rely solely

[18] Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2007. National Center for Health Statistics.

on cell phones may be different in important ways from other members of the same subgroups.[19]

These phenomena have several implications for the MATS 2007 survey estimates. Some types of adults may be under-represented in the survey sample, young adults may be even more under-represented, and certain types of young adults still more under-represented. While the raking step for creating the sample weights can adjust for some of this difference, it cannot adjust for all of it, particularly for those characteristics that are not associated with the characteristics used for the raking – age, gender, race, and education. To the extent that adults who have only cell phone service, and younger adults in general, are different from older adults in relation to their smoking behavior and other characteristics of interest to MATS, the statistics reported for the overall population may be affected by the under-representation of young adults. To the extent that younger adults in cell phone-only households are different from all young adults, the statistics reported for the young adult subgroup may be affected by the absence of the cell-phone only households in the MATS 2007 sampling frame.

The Blue Cross sample was likely to include some people with only cell phone service. If it was possible to locate such members and interview them, then there would be some representation of individuals from cell phone-only households in the MATS 2007 sample. Any corrective effect so produced is small and unknown. Their proportion in the final merged sample would be very small, and they would still not be a random sample of cell phone-only individuals in the general Minnesota population, since they were drawn from those who have Blue Cross insurance. The merging of the Blue Cross sample with the RDD sample may, at best, have slightly mitigated the cell phone-only undercoverage problem.

[19]It is important to emphasize that the issue is the percentage of individuals living in cell phone-only households. An individual may personally choose to rely exclusively on a cell phone, but if he or she is a member of a household with a traditional landline phone, that individual was covered by the MATS 2007 RDD frame and still could be included in the sample.

6.2.2.2 Measurement Error

Nonsampling errors in surveys may be attributed to a variety of sources, many of which fall under the type called measurement error. These sources of potential error may result from how the survey was designed, how respondents interpret questions, how able and willing respondents are to provide correct answers, and how accurately the answers are recorded and processed. MATS 2007 took several steps to minimize these types of errors. Important ones for MATS 2007, as described throughout this methodology report, include the careful and deliberate design of the questionnaire with review by multiple individuals and organizations; improvements to the clarity of the questions, balanced against possible effects on comparability of the responses across survey rounds; the use of a CATI system to administer the questionnaire and record responses; the internal testing of the CATI questionnaire; the pilot test of the instrument and survey procedures; the monitoring of the sample and of the collected data throughout data collection; and the thorough review of that data prior to finalizing the file for analysis.

The weighting process – especially the raking/post-stratification adjustments – partially corrects for bias due to minor discrepancies in the representativeness of the sample. During the weighting process, extensive diagnostic examination of the effects of the weighting design and of draft weights on the weighted estimates of demographics, smoking prevalence, and other characteristics further supported the "calibration" of the sample into closer conformity with the overall Minnesota population. Moreover, biases also may be present when people who are missed in the survey differ from those interviewed in ways other than the categories used in weighting. People who are missed in the survey include those missed because of the frame coverage issue or because sampled individuals did not respond to the survey. As with most surveys that rely on telephone interviewing, it is likely that some subgroups, such as specific racial and ethnic minority communities, are under-represented. The MATS 2007 sample design sought to correct for anticipated shortfalls, in two such groups, African Americans and young adults.

All of these considerations affect comparisons across different surveys or data sources. Although most of these limitations are inherent in all surveys, MATS 2007 made every effort to minimize these limitations.

Sources

Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2007. National Center for Health Statistics. Available from: http://www.cdc.gov/nchs/nhis.htm. May 13, 2008.

Rizzo L, Brick JM, Park I. A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys. Public Opinion Quarterly 2004 68(2):267-274.

Appendixes

A. MATS 2007 Questionnaire

B. MATS 2007 Questionnaire Skip Pattern Table

C. MATS 2007 Household Screener Questionnaire

D. MATS 2007 Letters

- D1-D3. RDD: Advance, Non-Response, Refusal
- D4-D6. Blue Cross: Advance, Non-Response, Refusal
- D7-D9. First Plan: Advance, Non-Response, Refusal

E. MATS 2007 Web Page Content

F. Significant Modifications of MATS 2003 Questionnaire Items for MATS 2007

G. Significant Modifications of MATS 2003 Skip Patterns for MATS 2007

NOTE: Appendixes A and F are available at www.mnadulttobaccosurvey.org. The remaining appendices and other reports referenced in this report can be requested by contacting Ann St. Claire at ClearWay Minnesota at astclaire@clearwaymn.org or by calling 952-767-1400.