Appendix A: Detailed Methodology

The national study contains the results of 6,139 interviews among adults in the U.S., of whom 1,247 were caregivers. A total of approximate 200 caregiver interviews each of African American, Hispanic, and Asian caregivers were obtained.

A companion study was conducted for AARP in seven states, resulting in approximately 400 caregiver interviews in each: California (from a total of n=2,040 interviews), Delaware (n=1,827), Illinois (n=2,072), Kansas (n=1,833), Ohio (n=1,864), Virginia (n=1,885), and Washington (n=1,979).

The sections below describe the basic research and sample designs for the study, the research procedures used to implement the project, the methodology used for estimating caregivers, weighting specifics, and descriptions of cross-tabulated variables. The response rate is included at the end of this section.

A. Research Design

The research design of this study is based on achieving several goals for NAC and AARP. One driving goal of the project is to estimate the number of caregivers in the U.S. population, as well as the number of caregiving households. Other aims include comparing caregiving roles and impacts across racial and ethnic groups and expanding on the understanding gained from the 1997 NAC/AARP caregiving survey.

In conjunction with NAC and AARP, Belden Russonello & Stewart (BRS) and Research/Strategy/Management, Inc. (RSM) designed a survey research project to achieve these goals. The project design included developing a methodology for calculating incidences, drafting and pre-testing a questionnaire, and devising a series of sample frames to achieve a nationally representative sample of caregivers, as well as over-samples of three racial/ethnic groups (African Americans, Hispanics, and Asian Americans).

B. Questionnaire Design

The questionnaire was drafted by BRS, NAC, AARP, and RSM, and drew from the 1997 NAC/AARP study, *Family Caregiving in the U.S.* The questionnaire has two components. The first comprises a screener used to a) identify caregiver(s) in the household, and b) obtain a host of demographic information from a randomly-selected respondent in all households contacted – regardless of the presence of caregiving in the household. As described below, these data are necessary to compile the base from which we estimate the proportion of caregivers.

The second part of the questionnaire reflects the substantive questions administered to all caregivers. For the state surveys, a shorter version of the questionnaire was used. See Appendix D for the CATI versions of the questionnaire.

After undergoing several iterative revisions, a final draft was programmed into a CATI system and pre-tested. BRS participated in training the interviewers, who were then monitored by members of the research team.

Based on the results of the pretest and a debriefing of the interviewers, changes were made in question wording and length. The questionnaire was then translated into Spanish for those respondents who preferred to be interviewed in Spanish. BRS back-coded the Spanish questionnaire, and input from AARP helped polish the final version. The interviews averaged 20 minutes in length for the national study, and 10 minutes for the state surveys.

C. Sample

1. Sample design

The research design for the national study called for completing about 1,200 interviews with caregivers, including samples of at least 200 caregiver interviews among African Americans, Hispanics, and Asian Americans. To achieve this result in the most cost-effective way, we used a combination of RDD, surname and targeted RDD sample frames. The supplemental samples (surname and targeted RDD) were implemented after the main RDD sample was finished and the number of remaining caregiver interviews in each subgroup needed to reach 200 interviews could be determined.

Soon after interviewing began with the supplemental samples, we noticed the productivity was much lower than anticipated (based on 1997 incidence reports for minorities in the NAC/AARP study) for the Asian surname, Hispanic surname, and targeted Hispanic samples. For example, 3% of contacts from the Asian surname sample produced an English-speaking caregiver. This low incidence raised questions about previously reported incidence among these minorities, was prohibitively expensive, and would take much longer to obtain the necessary interviews than time permitted.

Based on these factors and the dearth of other supplemental sample, the research team consulted with Knowledge Networks to draw on their panel of Americans to complete the remaining Asian and Hispanic interviews. Knowledge Networks has recruited the first online research panel that is designed to be representative of the entire U.S. population. The panel is recruited using high quality probability sampling techniques, and is not limited to current Web users or computer owners. Knowledge Networks selects households using random digit dialing (RDD) and provides selected households with free hardware and Internet access for participating in online surveys.[1]

RDD sample was used exclusively for the states surveys, obtained from Survey Sampling, Inc. (SSI).

2. Sample specifics for national study

The national RDD sample, also obtained from SSI, was stratified by geography to generate a set of telephone numbers proportionate to the population. The telephone numbers were then organized into replicates of 200 numbers each. SSI

---

[1] For more information, see www.knowledgenetworks.com

provided the targeted RDD samples, which were drawn from telephone exchanges in areas with higher than average (at least 30 percent density) concentrations of Hispanics and African Americans. SSI also provided randomly selected samples from a list of telephone numbers across the country linked to Hispanic and Asian surnames.

Knowledge Network's entire Asian sample was used for screening, and a random sample of their Hispanic panelists was drawn for the study. While panelists typically participate in online surveys, Knowledge Networks allowed us to screen for caregivers using an online screener, but administer the substantive questions to caregivers by telephone. Thus we were able to minimize any mode effects in caregiver results.

A total of 1,247 caregivers were obtained from the following samples:

- The main RDD sample produced 3,684 total interviews;
- Targeted samples produced 1008 total interviews; and
- Knowledge Network sample produced 1,447 total interviews.

From all of these interviews, we obtained interviews from 628 white, non-Hispanic caregivers, 200 African-American caregivers, 204 Hispanic caregivers, and 200 Asian-American caregivers.

3. Comparing data from different sample types for national study

Data from all samples were combined and weighted for final analysis. Prior to that, the results from each sample were first compared with Census and the main RDD sample, where appropriate, in effort to detect bias stemming from the different frames utilized in the study. Comparative analyses were conducted in two steps: 1) comparing sex, age, and race of all randomly-selected respondents; and 2) comparing sex, age, and race of all caregiver respondents. For example, all African-American randomly-selected respondents from the targeted RDD over-sample were compared with their counterparts in the main RDD sample. Additionally, demographics of the African American caregivers from the targeted sample were compared with those obtained from the main RDD. These comparisons indicted that the best representative sample would be a combination of all, weighted to known population or household parameters, depending on the level of analysis.

D. Field Methodology

Interviews from the main RDD national sample, the African American over-sample, and the state surveys were conducted September 5 through October 14, 2003. The Hispanic and Asian over-samples continued through December 22, 2003.

A five-call design was used for interviewing so that every number was dialed at least five times in an attempt to establish contact. Every soft refusal was followed by another attempt to convert the refusal into a completed interview.

When contact was established, the interviewer asked to speak with the person in the household aged 18 or older with the most recent birthday – our random selection process. In the case of the racial/ethnic over-samples, race was also determined for eligibility. The interviewer continued with the screener among all randomly-selected respondents, and proceeded to the substantive portion of the questionnaire if a caregiver was identified and agreed to participate.

For the Knowledge Networks panel, the screener portion of the questionnaire was administered to all of the Asian households on their panel, and a randomly selected sample of their Hispanic panelists. If a caregiver in the household was identified, the name, phone number, and a convenient time to call were obtained. If more than one caregiver was identified, Knowledge Networks randomly selected one to participate in the telephone interview. Those data were then forwarded to the field service to complete the caregiver interviews by telephone. All caregiver interviews were conducted by telephone instead of online to eliminate the chance of error due to mode effects. Data from all online screeners and telephone interviews among Hispanics and Asians were collected.

E. Definition of a Caregiver

For the purposes of this study, a caregiver was defined using two mechanisms: self report from a screening question, and verification through reporting assistance with at least one ADL or IADL.

The screening question used for the randomly selected respondent to identify the presence of a caregiver in the household is as follows:

> In the last 12 months, have you or anyone in your household provided unpaid care to a relative or friend 18 years or older to help them take care of themselves? Unpaid care may include help with personal needs or household chores. It might be managing a person's finances, arranging for outside services, or visiting regularly to see how they are doing. This person need not live with you. (IF YES: IS THAT YOU OR SOMEONE ELSE IN THE HOUSEHOLD?)

If a caregiver did not report providing help with at least one ADL or IADL in our series, the respondent was thanked for her/his participation and the interview was terminated. For all analytic purposes in this study, these self-reported caregivers are included as non-caregivers in the base of the 6,139 interviews.

F. Estimating Incidence of Caregiving

Estimating the number of caregivers and caregiving households for the national study involves two major components. The first component is the collection of all data necessary to make the estimates. This includes randomly selecting a respondent from all households contacted and obtaining a) demographic data on that individual; b) information about their household composition; c) the report of a caregiver(s) living in the household; and d) interviewing the caregiver.

The second component of estimating incidence involves weighting our large sample of data to a representative sample of the U.S. population and households, based on 2003 U.S. Census population projections and 2003 household projections from Harvard's Joint Center for Housing Studies (JCHS), which used U.S. Census 2000 household data to make their projections. From the weighted data we can estimate the proportion of caregivers and caregiving households in the U.S.

The same procedures were used for the state surveys, except we only make estimates for the population in each state – not households. Therefore, household data were not collected at this level, and we only created one weight based on 2003 population estimates in each state.

1. First Component: Collecting necessary data

The base from which to estimate caregiving in the U.S. – that is, a nationally representative sample of adults and households – was established through interviewing a randomly-selected respondent from each household contacted. Pertinent demographic data were collected on all of these initial, randomly-selected respondents (age, race, and gender). Furthermore, data on the household were gathered for the national study (size, family or non-family status of household members, age of householder, and race of householder).

All randomly-selected respondents were also asked whether or not there was a caregiver in the household. Three scenarios occurred based on the response from this question:

- If the randomly-selected respondent reported no caregiver in the household, the interview ended after obtaining the demographic data for the individual and the household;

- If the randomly-selected respondent was a caregiver, the full questionnaire was administered (only these initially selected respondents were used to determine population incidence to estimate the size of the caregiving population); and

- If the randomly-selected respondent was not a caregiver, but he/she identified someone in the household as a caregiver, the interviewer asked to speak with the caregiver after obtaining necessary demographic data from the randomly-selected respondent. The secondary respondent in a household was a caregiver by definition and was added to initial caregiver respondents to determine household incidence to estimate the number of households with caregivers. (Only one caregiver per household was interviewed.)

The same procedure was used for the states, with the exception of not obtaining household data. See Appendix D for the full CATI-formatted questionnaire with all instructions.

Interviewing continued until we obtained our desired number of caregivers.

2. Second Component: Weighting sample and making estimates

Two weights were computed for estimating in the national study: a "population weight" was used to estimate the proportion of caregivers in the *population,* and a "household weight" was applied to estimate the proportion of caregiving *households* in the US.

Estimating proportion of caregivers in the U.S. population

The population weight was a sex/age/race matrix based on the third quarter (October) 2003 US population projections by the US Census.

After applying the population weight to all data, we looked at the proportion of caregivers among the randomly-selected respondents in our sample. The result shows that caregivers comprise an estimated 21% of the U.S. adult population. The estimated number of caregivers in the US adult population is 44,443,800.

To estimate the incidence of caregivers helping recipients who are 50 or older, the same procedure was used, except we looked at the proportion of caregivers *giving care to 50+ recipients* among the randomly selected respondents in our sample. An estimated 16% of the adult population, or 33,861,900 adults, care for someone 50 or older.

Incidence of Caregivers in the U.S. Population[2]

	Incidence	X	2003 Census adult pop. estimate	=	Estimated number
Caregivers	21%		[211,637,100]		44,443,800
Caregivers of 50+	16%				33,861,900

The same procedures were used for estimating incidence in each of the seven states.

Incidence of Caregivers in Population by State Over-Sample

	Incidence	X	2003 Census adult pop. estimate	=	Estimated number
California	18%		23,840,900		4,291,400
Delaware	22%		590,700		130,000
Illinois	18%		9,033,000		1,625,900
Kansas	22%		2,026,600		445,900
Ohio	21%		8,603,500		1,806,700
Virginia	21%		5,477,400		1,150,300
Washington	19%		4,619,800		877,800

Estimating proportion of caregiving households in the U.S.:

Household estimates were calculated only for the national study. The household weight was based on the 2003 household projections derived from data supplied by JCHS, and constructed from a matrix of household type (family or non-family), age, and race of householder collected from all randomly-selected respondents.

Using the household weight, we calculated the proportion of all caregivers, including those identified by the randomly-selected respondent who was not him/herself a caregiver, among all households interviewed. We find 21% of U.S. households contain a caregiver(s). The estimated number of caregiving households in the U.S. is 22,901,800.

To estimate the incidence of households in which a caregiver(s) helps a recipient who is 50 or older, the same procedure was used, except we looked at the proportion of all caregivers *giving care to 50+ recipients* among the households interviewed.

---

[2] All estimates have been rounded to the nearest 100.

Incidence of Caregiving Households in the U.S.

	Incidence	X	2003 Census household estimate	=	Estimated number
Caregiving HH	21%		[109,056,000]		22,901,800
Caregiving HH of 50+	17%				18,539,500

Estimating proportion of caregivers within racial/ethnic subgroups

The same procedures for making national estimates were used in calculating incidence within racial/ethnic subgroups from the national study. To calculate the incidence of caregivers within a group, the base was either randomly-selected respondents within the designated race (for incidence in the population) or households of the designated race (for incidence of households). See Tables 1 and 2 in the report for percentages and estimated numbers of caregivers and caregiving households within each subgroup.

The following table reflects projections for 2003 total population and total households in the U.S. within each of the following racial/ethnic subgroups:

Population and Household U.S. Census Projections for 2003

	Adult Population N =	Households N =
Total	**211,637,100**	**109,056,000**
White, non-Hispanic	153,985,500	80,682,800
African American/Black	24,380,600	13,412,200
Hispanic	23,338,800	10,457,300
Asian-American	8,486,200	4,502,700

G. Weighting Substantive Results

As described above, a population weight and a household weight were created to estimate the incidence of caregiving among the population and in households in the U.S. For reporting the results of the national survey in this report among the 1,247 caregivers interviewed, the data weighted by households was used for two reasons. One is that the 1997 NAC/AARP caregiving study reported household-level data and where questions were similar we desired to make as valid a comparison as possible. The second reason is that the final caregiver sample was derived from identifying caregivers in the household even if they were not the initial, randomly-selected respondent. Hence, the sample is representative of caregivers in U.S. households.

For the state results, the same population weight was applied to the substantive results as was used to calculate incidence in the population of each state (each state had its own weight).

The margin of error for the sample of 6,139 interviews is $\pm$ 1.3. The margin of sampling error for our caregiver sample of 1,247 is conservatively estimated at $\pm$ 3.0 percentage points at the 95% level of confidence. Margin of sampling error for state samples, ranging from approximately n=1,800 to n=2,100 ranges from $\pm$ 2.2 to $\pm$ 2.3. For the substantive results in the states surveys, based on the n=400 interviews in each, the margin of error is $\pm$ 4.9 percentage points. Margin of error will be larger for subgroups within each sample.

H. Cross Tabulations

For the national study, cross tabulations were run on the total 1,247 caregivers, weighted by household. Cross tabulations were also run on each of the state samples of caregivers, with similar variables.

The following tables show key demographic and behavioral variables for the national study, followed by those used in the state studies.

Banner One: Descriptions of Variables

Heading	Labels	Description	Question Root
Gender	Male Female	Gender of caregiver	Recorded gender
Race	Wht	White, Non-Hispanic	Race, from screener
Age	18-34 35-49 50-64 65+	Age of caregiver	Age of caregiver, from screener
Education	<=HS SME CLG CLG+	High school or less Some college, technical college College degree or higher	Q94
Income	Less $30K $30-$49K $50-$100K $100K Plus	Less than $30,000 household inc. $30-$49K household income $50K-$100K household income $100K+ household income	Q95
Emp Care	Yes	Employed while caregiving	Either a) current caregiver *and* currently employed, or b) "yes" in Q40
	No	Not employed while caregiving	All others, not employed while caregiving
Rec. Age	18-49 50+	Care recipient is 18-49 years old Care recipient is 50+	Q3
Alzheim	Yes	Care recipient suffers from Alzheimer's or other mental confusion	Identified in Q11, or Q12
	No	Care recipient does not have Alzheimer's or other mental confusion	Q12
Distance	In HH < HR	Caregiver lives with recipient Caregiver lives less than an hour away from recipient	Q4
	1HR+	Caregiver lives one hour or more from recipient	

Banner Two: Descriptions of Variables

Heading	Labels	Description	Question Root
Care Status	Prim	Primary caregiver	Either "no", no one else provides unpaid help in Q31, or respondent provides most of the help in Q33
	Sec	Secondary caregiver	Someone else provides most of unpaid help or splits care in Q33
Hrs/Week	0-8	Less than 9 hours	Q27
	9-20	9-20 hours	
	21+	21 or more, including "constant care"	
Care Area	Urb	Caregiver lives in urban area	Q9
	Sub	Caregiver lives in suburban area	
	Rur	Caregiver lives in rural area	
Rec. Area	Urb	Care recipient lives in urban area	Q8 & Q9 (if care recipient lives in caregiver's household)
	Sub	Care recipient lives in suburban	
	Rur	Care recipient lives in rural area	
Health	Exc	Caregiver self reports "excellent" health	Q87
	Very Good	Reports "very good" or "good" health	
	Fair Poor	Reports "fair" or "poor" health	
Outside Resources	Yes	Reports using at least one of the outside/formal services	Q71-Q78
	No	Does not use any of these services	
Level of Burden	1	Each caregiver's score on two indices, an hours per week index and type of care index (see tables below) was summed, resulting in his/her assignment to one of seven levels (2, 3, 4, 5, 6, 7, or 8). As in 1996, the seven levels were collapsed into five, with Level 1 being the least intense level of caregiving, and Level 5 being the most intense.	Q27 and Q14-Q26
	2		
	3		
	4		
	5		

Level of Burden is the sum of the following two indices:

Hours of Care Index (Q27)		Type of Care Index (Q14-Q26)	
0-8 hours =	1	1 IADL / 0 ADLs =	1
9-20 hours =	2	2+ IADLs / 0 ADLs =	2
21-40 hours =	3	1 ADL (with or without IADLs) =	3
41+ or "constant care" =	4	2+ ADLs (with or without IADLs) =	4

Level of Burden:

	Combined score
Level 1	2, 3
Level 2	4
Level 3	5
Level 4	6, 7
Level 5	8

States' Banner: Descriptions of Variables

Heading	Labels	Description	Question Root
Gender	Male Female	Gender of caregiver	Recorded gender
Age	<50 50+	Age of caregiver	Age of caregiver, from screener
Education	<=HS Some Clg Clg+	High school or less Some college, technical college College degree or higher	Q94
Income	Less $50K $50K Plus	Less than $50,000 household inc. $50K+ household income	Q95
Emp Care	Yes No	Employed while caregiving Not employed while caregiving	Either a) current caregiver *and* currently employed, or b) "yes" in Q40 All others, not employed while caregiving
Rec. Age	18-49 50+	Care recipient is 18-49 years old Care recipient is 50+	Q3
Rec. Area	URB SUB RUR	Care recipient lives in urban area Care recipient lives in suburban area Care recipient lives in rural area	Q8
Out Res.	Yes No	Reports using at least one of the outside/formal services Does not use any of these services	Q71-Q77
Choice Care	Yes No	Had choice in taking on responsibility of caregiving Did not have choice in taking on responsibility of caregiving	Q59
Type of Care	No ADLs 1+ ADLs	No ADLs/1 or more IADLs 1 or more ADLs (with or without IADLs)	Q14-Q26

I. Regression Analysis

Multiple linear regression investigates the extent to which independent variables (*e.g.*, age, income, race, sex) influence a dependent variable(s) (*e.g.,* hours per week providing care). Regression also helps determine whether a statistically significant influence of an independent variable on a dependent variable indicates a real or superficial relationship by controlling for other independent variables.

Regression analysis was conducted to help determine what factors influence the emotional stress of caregivers, as indicated by a five-point scale in Q57. The following independent variables were used in the model:

Gender
Caregiver age
Education
Income
Distance from recipient
Caregiver status (primary or secondary)
Self-reported health
Caring for someone with Alzheimer's or dementia
Feeling a choice in taking on caregiving responsibilities
Level of Burden Index

Regressions were also run against feeling the caregiving situation has worsened one's health (Q88). The following independent variables were used:

Gender
Caregiver age
Education
Income
Distance from recipient
Caregiver status (primary or secondary)
Caring for someone with Alzheimer's or dementia
Feeling a choice in taking on caregiving responsibilities
Level of Burden Index

J. Response Rate

The response and refusal rates reported below are based on the Standard Definitions of the American Association for Public Opinion Research (AAPOR) for the final dispositions of cases codes and outcome rates for RDD telephone surveys.[3] Information regarding these calculations can be found on AAPOR's website (**www.aapor.org**).

The overall response rate for the national study is 19%, based on AAPOR's Response Rate 3 formula: $I/(I+P) + (R+NC+O) + e(UH+UO)$.[4] The response rate is higher for the RDD portion of the study (27%), and lower for each of the over-samples conducted from surname and targeted samples (ranges from 5% to 17%).

---

[3] The American Association for Public Opinion Research. 2000. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, Michigan: AAPOR.
[4] Interviews from Knowledge Networks are not included in the response rate calculation because dispositions for online screening are not comparable.