Tracing Women Over Half a Century : Strategies to Locate Subjects Lost to Follow-Up in a Longitudinal Health Study

Jan Root, Ken R. Smith, Elizabeth A. Whelan, Dale Sandler and Ann M. Voda

Research on Aging 1994 16: 375
DOI: 10.1177/0164027594164002

Tracing Women Over Half a Century

Strategies to Locate Subjects Lost to Follow-Up in a Longitudinal Health Study

JAN ROOT
Utah Health Information Network
KEN R. SMITH
University of Utah
ELIZABETH A. WHELAN
National Institute for Occupational Safety and Health
DALE SANDLER
National Institute of Environmental Health Sciences
ANN M. VODA
University of Utah

Cohort studies typically require repeated contacts with the same study participants over many years. In large-scale follow-up studies of women, it can be difficult to locate and maintain contact with participants because women are more likely to experience name changes through marriage and divorce and may be less visible in public records, particularly contemporary elderly women. This article reports on the methods used in 1990 and 1991 to trace 998 participants of the Women's Health Study (WHS) who were enrolled during the 1930s in the Menstruation and Reproductive History study, a longitudinal study on menstruation and reproduction. Some of these women had been lost to follow-up for nearly 50 years. This article reviews the strategies used to locate the WHS participants.

In large-scale follow-up studies of women, it can be difficult to locate and maintain contact with participants because women are more likely to experience name changes through marriage and divorce and may be less visible in public records, particularly contemporary, elderly women. This article reports on the methods used in 1990 and 1991 to trace 998 participants of the Women's Health Study who were enrolled during the 1930s in the Menstruation and Reproductive History study.

RESEARCH ON AGING, Vol. 16 No. 4, December 1994 375-388

Some of the 1930s enrollees had been lost to follow-up for nearly 50 years.

Cohort Description

The Tremin Trust Research Program (TTRC), originally known as the Menstruation and Reproductive History (MRH) research program, began at the University of Minnesota (UM) in 1935 under the direction of Alan Treloar (Voda, Morgan, Root, and Smith 1991). The purpose of the MRH research program was to describe characteristics of the human menstrual cycle (Treloar, Boynton, Behn, and Brown 1967). More than 2,600 women were enrolled in the MRH program between 1935 and 1939 (Cohort 1). An additional 1,600 women were enrolled between 1961 and 1963 (Cohort 2). At the time of enrollment, most participants in Cohorts 1 and 2 were entering freshmen at the University of Minnesota. The MRH research program was moved to the University of Utah in 1984 and was renamed the Tremin Trust ("Tre" for Treloar and "Min" for Minnesota) Research Program (TTRP). Limited funding of the study over the years restricted efforts to trace participants who dropped out of the study.

The Women's Health Study (WHS) began in 1986 and was designed to evaluate the effects of menstrual and reproductive characteristics on several health outcomes in later life based on 998 women from Cohort 1. Time since last contact with the WHS women ranged from 1 to 48 years (Table 1).

One of the main tasks of the WHS involved identifying current addresses for all eligible participants and administering a mailed questionnaire. When a participant was identified as deceased or incapacitated, an attempt was made to locate a knowledgeable proxy respondent for that person.

Resources Used

The strategies used to trace study participants and/or proxy respondents included widely available resources as well as infor-

TABLE 1
Years Since Last Contact With
Women's Health Study Participants as of April 1989

Years Since Last Contact	N	*Percentage*
40-48	115	11.5
30-39	86	8.6
20-29	58	5.8
10-19	175	17.5
1-9	564	56.5
Total	998	100.0

mation unique to this cohort. These resources are summarized in Table 2.

Often two or more tracing strategies were used concurrently for the same participant, and it was not uncommon for a participant to be located through more than one strategy. The WHS was not designed to evaluate the relative effectiveness of the various tracing strategies. However, to the extent possible, the proportion of participants traced using a given strategy and the proportion found by that method are given in Table 2. Most of the strategies employed in the WHS could be used in other follow-up studies of adults, male or female. Strategies were grouped into two areas, those used to locate participants believed to be alive and/or proxy respondents and those used to identify the deceased.

The term *match* is defined as a *tentative* identification of a participant's whereabouts or vital status. The term *find* signifies that tentative information about the participant led us to the living participant or proxy or to records that confirmed a death. A find or match could have occurred more than once based on two different tracing strategies.

STRATEGIES SPECIFIC TO THE WHS

MRH/Tremin Trust Research Program Files

The first strategy involved a review of each participant's MRH file for information that would help locate eligible participants. This information included full name, maiden name, all married names and

TABLE 2
Summary of Search Strategies Used in the Women's Health Study

Strategy[a] *(information used for tracing)*	*Number of Persons Searched*	*Total Found/ Deceased*	*Approximate Fee*
University of Minnesota Foundation (Alumni) (name, address, DOB sent)	681	147/18	$170
Wisconsin driver's license records (name, DOB, address, SSN)	388	10	free[b]
Minnesota marriage license (name, year of marriage)	400	2	free[b]
University of Minnesota yearbook/sororities (request for copies of graduating class, 1935-1945; request address information)	650	49	approx. $100[b]
Health Care Financing Administration (name, state of residency, sex, DOB, SSN)	998	333/44	free[c]
U.S. Post Office (name, address)	111	17	free[c]
Telephone search (name, address, other family names and addresses, hometown, other known addresses)	not known	164	varies by state

California driver's license records (name, DOB, address, SSN)	87	9	free
Minnesota driver's license records (name, DOB, address, SSN)	97	1	$4 per search
Minnesota motor vehicle records (name, DOB, address, SSN)	66	16	$1 per search
Credit bureau (name, DOB, address, SSN)	231	53	<$1 per search
Social Security Administration Master Beneficiary Record (computer file, name, state of residency/death, DOB, DOD, SSN)	998	65	free
National Death Index (computer file, name, father's last name, state of residency/death, DOB, DOD, SSN)	723	71	$350 for an initial search of sample, $0.30 per record charge, $1.00 keying charge per record; $8 to obtain copies of death certificates from state vital statistics offices[d]

NOTE: DOB = date of birth; DOD = date of death; SSN = Social Security number.
a. See text for an explanation of these strategies and the protocol used with each in the Women's Health Study.
b. Labor intensive.
c. Available to government agencies.
d. National Center for Health Statistics (1990).

dates of marriages, Social Security number (SSN), date of birth (DOB), all current and previous addresses and dates of addresses, any information on other family members or friends, years enrolled in the MRH program, and, if the participant was known to be deceased, the date of death (DOD) and place of death. The SSNs were unavailable for most participants. The SSN system did not exist at the inception of the MRH program, and its potential for tracing lost participants was not appreciated until the mid-1970s. However, a participant's maiden name was almost always known, and it was the maiden name that became a key element in establishing matches.

University of Minnesota Foundation

The University of Minnesota Foundation (UMF) is an organization that maintains contact with its alumni. Nearly all of the eligible WHS participants were university students although some did not graduate. The UMF mailing list is updated as UMF is informed of changes in address or vital status of UM alumni, but no systematic attempt is made to locate alumni who are "lost," a common practice among alumni associations. UMF maintains records only on students who have graduated. For tracing purposes, UMF constituted an efficient and inexpensive tracing resource. Other studies based on college populations will also find alumni associations to be an invaluable resource.

University of Minnesota Yearbooks/Sororities

It was common during the 1930s for female students to belong to a sorority. Yearbooks during that time period contained information on sorority membership. Some yearbooks also contained information on a participant's hometown, data that were used in subsequent search strategies. The executive directors of all 18 sororities were contacted, and 9 provided addresses or vital statuses for participants.

This strategy is attractive in that college yearbooks are accessible and inexpensive. The usefulness of any yearbook for tracing purposes largely depends on the information it contains. The sororities contacted expressed varying degrees of concern regarding the confidentiality of mailing lists, ranging from an open policy to being extremely

guarded. Sororities unwilling to divulge information agreed to forward our inquiries to participants.

STATE-LEVEL SEARCH STRATEGIES

Minnesota Marriage License Bureau

Studies of women must consider the possibility that name changes will occur, particularly for women lost to follow-up for an extended period of time. Marriage records for those participants who were last known to reside in Minnesota were searched for name changes that might have occurred after participants left the study. Securing the most recent name of the participants was crucial for using subsequent search strategies. Whereas it was possible to search marriage records for all states, it was most efficient to check for marriages in Minnesota.

For this cohort of women, it is possible that names did not change very often because divorces and remarriages were not common. Tracing more recent cohorts of women may require more attention to name changes given the increasing rates of divorce over the past two decades. For the WHS, this strategy was not productive. The major logistical problem was that *complete* marriage records in Minnesota are maintained only by county.

Driver's License Records/Motor Vehicle Records

The reason for using driver's license information for tracing is that this source generally contains recent residential information and is relatively accessible. Searches of driver's license and/or motor vehicle records were conducted in Minnesota as well as in the two states containing a large number of MRH women, California and Wisconsin. The tracing work was conducted by employees of the motor vehicle or driver's license offices.

We had a low rate of matching names to driver's license and/or motor vehicle registration records because a match was identified only when there was an *exact* match between the information supplied by the study and that on record in the three states searched. Other states may have more flexibility in their matching procedures.

It was advantageous to search both driver's license and motor vehicle records because they constitute different databases. Driver's licenses are renewed every five years in most states. Thus the address information obtained from this method may not be current. Motor vehicle record searches are also limited in that they contain information only for persons who have registered a car (i.e., husband's name). However, most motor vehicle records must be renewed yearly and, hence, if a person was found through this strategy, the address information was likely to be current. The difference in the match rates between the driver's license (1%) and motor vehicle records (24%) searches in Minnesota may have been due to differences in how the searches were implemented by the state agencies involved.

Telephone Directories

Telephone listings are updated annually and, generally, will contain recent information on addresses and numbers. Three distinct strategies were used: (a) a search of published telephone directories of major metropolitan areas in the United States; (b) use of Locator Plus, a computerized telephone listing of 14 Western states including Minnesota; and (c) a search by area code in California, the state to which most participants moved. For telephone directories and Locator Plus, the search began in the geographical and metropolitan areas closest to the last known address of the participants. Hometown names, obtained from UM yearbooks, were also included in the geographical areas in an attempt to locate participants' relatives who might be living there. Locator Plus was more convenient to use than were the telephone directories in that all persons with the same last name as the participant from the 14 states were searched together via computer. The search of California area codes involved submitting a list of participants' names and their last known addresses to an operator, who then searched their directories.

Searching names through telephone directories is labor intensive: a large number of possible matches are generated, most of which can be eliminated only by contacting each person on the list. For this study, more than 10,000 calls were placed during a 6-month period. Often a find was the result of perseverance on the part of the caller. Sometimes

100 calls were made to locate a single individual. Locator Plus was advantageous because all individuals with a particular last name could be examined together via computer, and it was not limited to major metropolitan areas.

The difficulties that arose using published telephone directories were that (a) access was limited to major metropolitan phone directories and (b) telephone directories are extremely cumbersome and time consuming to search, thereby increasing the chance of making matching errors. The major limitation of Locator Plus was that it did not include California, the state to which most participants had migrated.

NATIONAL-LEVEL SEARCH STRATEGIES

Health Care Financing Administration

Because surviving Cohort 1 participants would now be over 70 years of age, Medicare beneficiary records were considered a useful source of tracing information. Health Care Financing Administration (HCFA) monitors all health care claims through Medicare. HCFA records can be used to search participants who are either alive or recently deceased. When a Medicare beneficiary dies, HCFA records occasionally provide the name and address of the surviving beneficiary who may be available to act as a proxy respondent. Access to HCFA records is restricted to government agencies for approved research projects under strict guidelines for confidentiality and maintenance of records. The WHS was able to use HCFA data because the study was carried out in collaboration with the National Institute of Environment Health Sciences (NIEHS).

The HCFA search did not use the New York State Information and Identification System (NYSIIS). NYSIIS matches names phonetically in addition to names that are spelled identically. For example, NYSIIS can match Peterson with Petersen or Pedersen, matches that would be missed in exact spelling matches. HCFA matches on last names based on exact spellings. We could not, therefore, investigate potential matches for names that were similar but possibly misspelled, thereby underestimating the proportion of women in this study who were in the HCFA database.

The absence of SSN and accurate DOB for all participants reduced the efficiency of this strategy. An advantage of this strategy was that we could work directly with HCFA staff to adjust the matching criteria.

U.S. Post Office

It was not uncommon for participants to fail to inform the MRH/TTRP of their changes of address. U.S. Post Offices are required to forward mail, if requested, for a minimum of 3 months. The post office will verify current addresses or provide forwarding information to government agencies. This information may be available to others for a nominal fee. U.S. Post Offices serving smaller communities generally responded faster than did those in larger urban areas and kept information on forwarding addresses for more than 3 months.

Credit Bureau

Near the end of the study, a nationwide credit bureau was enlisted to search credit records for address information on participants who were still lost to follow-up. Whereas this company would search, with varying degrees of intensity, for individuals lost to follow-up in longitudinal studies, at our request it searched only two credit record databases because that was the most cost-effective methodology. The credit bureau suggested more aggressive methods, but such methods were considered too invasive for the WHS. If this method had been implemented earlier in the tracing process, the information generated might have been more useful, but it was considered too costly for our study.

STRATEGIES TO IDENTIFY DECEASED PARTICIPANTS

Social Security Administration Master Beneficiary Files

The Social Security Administration (SSA) maintains a Master Beneficiary Record (MBR) on all persons who died after 1962 and who maintained a Social Security account. An MBR is generated when a Social Security death beneficiary claim is made and/or benefits are

terminated due to death. SSA-MBR files identified only deceased participants.

A computerized list was submitted to SSA of all participants detailing first and last name, DOB, state of last residence, SSN, and, if deceased, state of death and DOD. The SSA-MBR search generated approximately 10,000 possible matches that were weighted according to their degree of matching. A case-by-case evaluation was made of the possible matches.

The lack of SSNs dramatically increased the number of possible matches. Accordingly, sorting through the returned possible matches, identifying and evaluating potential true matches, and then verifying each possibility through receipt of a death certificate was time consuming. SSA-MBR records include deaths from 1962 up to the present. The extensiveness of the years covered gives the SSA-MBR an advantage over the National Death Index (NDI), for which coverage begins in 1979. A further advantage of this strategy was that the SSA-MBR computerized search used NYSIIS.

National Death Index

The NDI is a computerized registry of all deaths occurring in the United States since 1979. The age of Cohort 1 participants was such that the majority of deaths were likely to have occurred during the last 10 years, thereby making NDI a potentially useful resource for this study.

The lack of SSNs and accurate DOBs reduced the efficiency of the NDI search. In lieu of an SSN, the matching element that was most important was father's last name (i.e., participant's maiden name).

Death Certificates

Copies of death certificates were sought from the appropriate state vital statistics office for participants who were believed to be deceased. This process was straightforward if the DOD and state of death were known in advance, as was the case for deaths identified through HCFA, SSA, or NDI. However, for family-reported deaths, deaths that were noted in the Tremin Trust files, or SSA- or NDI-reported deaths

in which the death certificate name or DOB was not an exact match, it was often more difficult to obtain copies of the correct death certificates.

Death certificate requests were made to vital statistics offices in 41 states. Each state has its own procedures to handle requests for death certificates. Each state charged various amounts, but the cost averaged $8.00 per search regardless of whether a match occurred. A complete listing of all state vital statistics office addresses can be obtained from NDI staff.

Death certificates were used to locate proxy respondents for deceased participants when allowed by state law. Information that was used to locate proxy informants included informant's name and address, cemetery/crematorium name, and funeral home name. Cemeteries were useful because they occasionally maintained family plots or the owner of the cemetery knew the family. As a result, the cemetery management occasionally had recent information on surviving relatives of deceased participants. Funeral homes also maintain information on families and are sometimes willing to share that information.

The procedures to request death certificates from states' vital statistics offices are variable. Some states require only the identification of the deceased person, payment of a fee, and an address to which to send the death certificate. Other states have more elaborate application procedures.

Because of the current budget deficits facing many state governments, some vital statistics offices are understaffed and hence may have difficulty meeting requests for death certificates. Obtaining death certificates for very recent deaths was not generally possible at the state level.

Discussion

In beginning the WHS, there was concern that the MRH women lost to follow-up would be extremely difficult to trace. Nonetheless, after implementing the strategies described, 95% of the 998 participants were successfully traced (dead or alive). Of these, 170 were confirmed deaths (either family report or death certificate) and 773

were found alive. Only 55 participants were not found. In addition to finding 95% of the participants, more than 160 proxy respondents were identified who completed questionnaires. Investigators who study the next several waves of older cohorts may find that they may have even greater success in tracing their subjects given that successive cohorts are more likely to have accurate SSNs available for tracing purposes.

Throughout the study, the exact criteria for what constituted a *find* remained somewhat flexible. The best matches were those based on telephone conversations with the respondent or proxy when it was possible to confirm historical data recorded in the MRH archives.

As with other studies that involve tracing lost participants, success came through exploiting the unique qualities of the cohort (Boice 1978; Cadell, Cantor, Parkes, and Barbano 1987). In this case, the centralization of Cohort 1 in Minnesota and the fact that the participants were nearly all students at the time of enrollment greatly assisted efforts to locate the participants. In addition, the use of nationally based strategies such as SSA, HCFA, and NDI increased the match rate significantly.

Recommendations to other researchers attempting to locate lost members of a longstanding cohort include (a) exploit unique tracing information about the cohort, (b) begin tracing with cost-effective strategies as well as those that may require processing time, and (c) implement tracing strategies early in the study that are free of charge. All tracing procedures that are implemented by researchers need to take precautions to protect the subjects from potentially harmful breeches of confidentiality and trust. Finally, researchers involved in tracing subjects should recognize that the advantages and disadvantages of the strategies described here can change as new policies are implemented that guide access to confidential data.

REFERENCES

Bean, J. A., J. D. Leeper, R. B. Wallace, B. M. Sherman and H. Jagger. 1979. "Variations in the Reporting of Menstrual Histories." *American Journal of Epidemiology* 109:181-5.

Boice, J. D., Jr. 1978. "Follow-up Methods to Trace Women for Pulmonary Tuberculosis, 1930-1954." *American Journal of Epidemiology* 107:127-39.

Cadell, D. M., D. Cantor, K. Parkes and H. E. Barbano. 1987. "NHANES I Epidemiologic Follow-Up Study: Tracing Experiences." *Proceedings of the 1987 Public Health Conference on Records and Statistics* (DHHS Publication No. [PHS] 88-1214), pp. 253-8.

National Center for Health Statistics. 1990. *National Death Index User's Manual* (DHHS Publication No. [PHS] 90-1148).

Treloar, A. E., R. E. Boynton, B. G. Behn and B. W. Brown. 1967. "Variation of the Human Menstrual Cycle through Reproductive Life." *International Journal of Fertility* 12:77-126.

Voda, A. M., J. Morgan, J. Root and K. R. Smith. 1991. "The Tremin Trust: An Intergenerational Research Program on Events Associated with Women's Menstrual and Reproductive Lives." Pp. 5-18 in *Menstruation, Health, and Illness*, edited by D. L. Taylor and N. F. Woods. New York: Hemisphere.