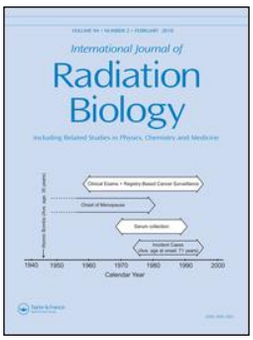

International Journal of Radiation Biology

ISSN: 0955-3002 (Print) 1362-3095 (Online) Journal homepage: http://tandfonline.com/loi/irab20

Obtaining vital status and cause of death on a million persons

Michael T. Mumma, Sarah S. Cohen, Jennifer L. Sirko, Elizabeth D. Ellis & John D. Boice Jr.

To cite this article: Michael T. Mumma, Sarah S. Cohen, Jennifer L. Sirko, Elizabeth D. Ellis & John D. Boice Jr. (2018): Obtaining vital status and cause of death on a million persons, International Journal of Radiation Biology, DOI: 10.1080/09553002.2018.1539884

To link to this article: https://doi.org/10.1080/09553002.2018.1539884

Accepted author version posted online: 09 Nov 2018.
Published online: 04 Jan 2019.

INTERNATIONAL JOURNAL OF RADIATION BIOLOGY
https://doi.org/10.1080/09553002.2018.1539884

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Obtaining vital status and cause of death on a million persons

Michael T. Mumma[a], Sarah S. Cohen[b], Jennifer L. Sirko[a], Elizabeth D. Ellis[c] and John D. Boice, Jr.[d,e]

[a]International Epidemiology Institute, Rockville, MD, USA; [b]EpidStat Institute, Ann Arbor, MI, USA; [c]Center for Epidemiologic Research, Oak Ridge Associated Universities, Oak Ridge, TN, USA; [d]National Council on Radiation Protection and Measurements, Bethesda, MD, USA; [e]Department of Medicine, Division of Epidemiology, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA

ABSTRACT

Purpose: To present the methodology used to determine vital status and obtain cause of death (COD) within the Million Person Study of Low-Dose Health Effects (MPS). Data sources and vital status tracing techniques used to obtain vital status and COD for six (n = 424,238 subjects) of the ~20+ cohorts under study are described.

Methods and materials: A multistage approach using multiple sources of vital status information was used to determine vital status (or 'trace') study participants from as early as 1940 to the present. Mortality records from state departments of vital statistics and the Social Security Administration Death Master File (SSA-DMF) were matched to study participants by Social Security Number (SSN), full name, date of birth (DOB), and/or sex using deterministic and probabilistic algorithms. The National Death Index (NDI) and SSA Service for Epidemiological Researchers (SSA-SER) were used to obtain COD (after 1978) and verification of alive status, respectively. Online public records and ancestry services, death certificates, and specialized mortality sources were also utilized.

Results: For the MPS cohorts traced to date (nuclear power plant workers, industrial radiographers, atomic veterans, and workers at Rocketdyne/Atomics International, Mound nuclear facility, and Mallinckrodt Chemical Works), vital status was confirmed for over 90% of all study subjects in all but one cohort (88%). The ascertainment of COD was over 96% for all cohorts.

Conclusions: A hallmark of a high-quality epidemiologic cohort mortality study is a low percentage of subjects with unknown vital status and a low percentage of deaths without a COD. The sources and methods used for vital status tracing and COD determination for the MPS have been successful and should be useful for other investigators tracing large, historic study populations. Some of the approaches would be applicable for use in all cohort studies using regional-specific mortality data or modifications to the approach.

ARTICLE HISTORY
Received 23 August 2018
Revised 5 October 2018
Accepted 10 October 2018

KEYWORDS
Vital status; cause of death; cohort study; social security; National Death Index

Introduction

In 2005, Buchanich et al. published a two-stage approach to vital status ascertainment, where stage 1 was submission of the study population to the Social Security Administration Service to Epidemiological Researchers (SSA-SER) and stage 2 was submission to the National Death Index (NDI). This approach was developed to obtain vital status in a cost-effective manner given the discovery of the incompleteness of the Social Security Administration Death Master File (SSA-DMF). Since then the incompleteness of the SSA-DMF has increased substantially, particularly since the passage of legislation in 2011 restricting the release of death records provided to the SSA from states.

The tracing process developed for the Million Person Study of Low-Dose Health Effects (MPS) expands upon the Buchanich et al. (2005) approach to efficiently obtain vital status for numerous study populations over nearly 80 years, and the methodology is detailed in this manuscript. Briefly, before going to SSA-SER, the MPS population was matched to a collection of in-house mortality data which included the Social Security Administration Death Master File (SSA-DMF) and mortality files assembled from 33 states. This matching is done using both custom algorithms developed by MPS collaborators with SAS software (SAS Institute Inc., Cary, NC) and the Centers for Disease Control and Prevention LinkPlus program, which does not require exact matches on all variables (Campbell et al. 2008). Unmatched participants are then submitted to the SSA-SER which may provide additional dates of death, but more importantly can provide a verification of alive status. At this stage, study participants known to have died in 1979 or later with cause of death (COD) still unknown or where vital status remains unknown are sent to NDI (which began in 1979). If study follow-up starts before 1979, an additional stage to obtain COD for deaths before 1979 must be completed (if COD was not already obtained from state mortality files). In the MPS, death certificates were obtained from a database of scanned death certificates, state departments of health, or online resources (e.g. Ancestry.com). During this multi-stage and iterative process,

CONTACT John D. Boice john.boice@vanderbilt.edu National Council on Radiation Protection and Measurements, 7910 Woodmont Avenue, Suite 400, Bethesda 20814, MD, USA

information from public record data providers are incorporated to confirm and/or correct key identifying information vital for successful matching and to validate a 'presumed alive' status for the small percentage of participants not verified as alive or deceased. Further details of the approach developed for the MPS are described in the Methods section.

Methods

Mortality data files

For over 40 years, MPS investigators have been assembling and consolidating mortality records, including over 57 million death records from 33 state departments of health and/or vital statistics (SDOH). Files from 21 of these 33 states contained causes of death, and 20 contained date of death only (8 states provided both over different time periods). The availability, years of coverage, cost and permissions required to obtain these files have varied widely among the states, with some states providing death information as early as the 1940s. The data items included in the files also vary. Most contain name, Social Security Number (SSN), sex, and date of birth (DOB), while some state files provide additional information useful in epidemiological investigations such as race and contributing causes of death. In addition, MPS investigators have obtained the SSA Death Master File (SSA DMF), containing over 86 million death records (https://www.ssdmf.com). Jones and Vawdrey (2015) provide an excellent description of the SSA DMF, its limitations, and how mortality data are collected in the United States. MPS investigators combined the state mortality files and SSA DMF into a single integrated mortality database (IMD) to facilitate the matching process.

Department of energy death certificates

In addition to the IMD, a collection of over 80,000 death certificates (DC) for workers employed at facilities in a Department of Energy (DOE) complex (or its predecessor) have been retrieved from SDOH and from Benefits Offices at the DOE facilities. Most of these DCs are for deaths that occurred prior to 1979. States in which DOE facilities were located, states surrounding those of employment, and states where workers were likely to retire were systematically targeted for obtaining DC. Standard algorithms were used to verify that the identifying and demographic data for the worker matched that information on the DC. Death certificates were scanned into electronic format (PDF) and indexed by name and/or SSN.

Governmental vital status services

Social security administration service to epidemiological researchers (SSA-SER)

The Social Security Act directs the SSA to provide epidemiological researchers information on whether study subjects are shown on SSA records to be alive or deceased (https://www.ssa.gov/policy/about/epidemiology.html) (Social Security Administration 2018). Requests to link cohort data to the SSA-SER must be approved through an application process and the study must be judged to 'contribute to a national health interest'. After approval is granted, SSN, full name, fullDOB, and sex of study participants are submitted for linkage.

The return file from SSA-SER places participants into one of three categories: dead, alive, or unknown. Death matches include a date of death and may also provide the state where a SSA benefit claim was filed or state of residence at death. Alive matches indicate the individual is living based on SSA program records such as Internal Revenue System, Medicare, or other benefits data. Unknown status indicates SSA has no record of death, insufficient data to presume the subject is alive, or there is missing or incorrect matching information. In addition, like the SSA DMF, the SSA-SER does not provide information on deaths that were provided from some states. Instead, participants where SSA was provided death information from certain states are categorized as 'unknown' vital status. This increasing proportion of 'unknown' vital status places additional importance on subsequent steps to completely ascertain vital status.

National death index

The NDI (https://www.cdc.gov/nchs/data/factsheets/factsheet_ndi.htm) is a centralized database of death record information compiled from state vital statistics offices and established as a resource to aid epidemiologists and other health and medical investigators with mortality ascertainment (Patterson and Bilgrad 1986). The NDI Plus service provides underlying and contributing causes of death coded according to the International Classification of Diseases (ICD) edition in place at the time of death (National Center for Health Statistics 2018). The NDI is an excellent source for identifying deaths occurring within the United States (Cowper et al. 2002). However, it is only available for deaths occurring in 1979 or later and, depending on the characteristics of the population under study, its completeness ranges between about 90–100% (Acquavella et al. 1986; Boyle and Decouflé 1990). Thus, multiple sources are important for complete ascertainment of deaths even after 1978 (Curb et al. 1985; Schall et al. 2001; Wojcik et al. 2010). The NDI can also be cost-prohibitive (Hermansen et al. 2009), particularly when COD is needed in addition to fact of death. Linking data to the NDI also starts with an application process. Once approved, records are sent to NDI in two groups: known deaths (searching NDI for COD only) and unknown deaths (searching NDI for both fact and COD). The cost for searches varies based on whether the death is known or unknown and the number of years searched. To reduce searching fees, particularly for large cohorts, it is helpful to determine the last date the participant was known to be alive and to search NDI only after this date. For occupational studies such as MPS, this is often the date of last employment or last occupational exposure monitoring. NDI returns linkages with a score indicating the quality of the match. We utilize an approach similar to Fillenbaum et al. (2009) where matches on SSN, last name, and one additional unique field (e.g. DOB or middle name) are accepted as matches without further review. The remaining NDI-provided

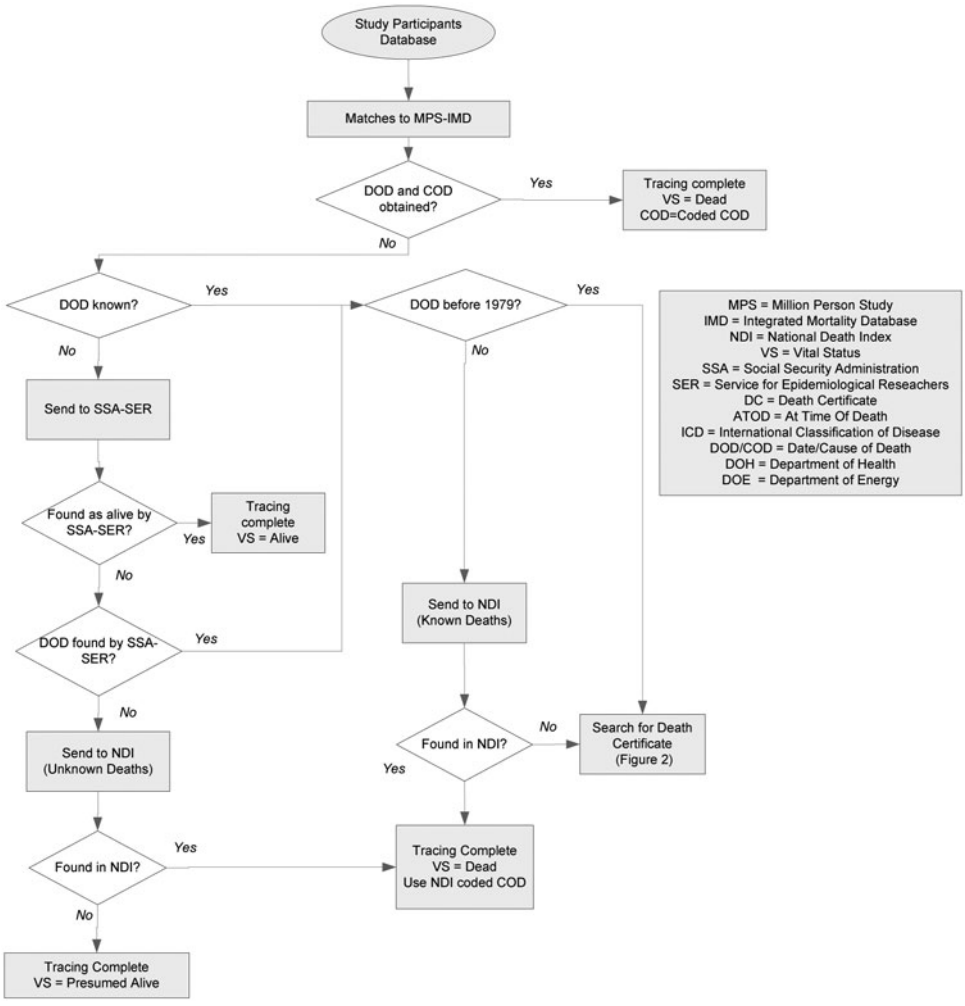

Figure 1. Generalized process to obtain vital status and cause of death utilized in the Million Person Study.

matches are evaluated on a case-by-case basis which can be time consuming.

Special sources

For certain occupational groups, special data sources were available to counter challenges in tracing and the lesser mortality coverage prior to 1979. For example, for military personnel who participated in weapons testing between 1945 and 1962, the Department of Veterans Affairs Beneficiary Identification Records Locator Service (BIRLS) allowed linkages based on the military ID number, particularly helpful as SSN was often not available or widely utilized in these early years (Page et al. 1996; Boyko et al. 2000; Maynard and Chapko 2004).

Vital status tracing process

The flowchart in Figure 1 shows a generalized, multistage approach to the MPS tracing process. The first stage is to link study participants to the combined state mortality and SSDMF mortality files (i.e. the IMD). The first match is conducted using customized SAS software code to obtain 'perfect' or highly likely matches. For example, a match on SSN, first and last name, gender, and DOB would be considered a perfect match. A match on SSN, last name, and one additional piece of information such as DOB or middle name or gender, would be considered as highly likely match. A second match is then conducted using the Centers for Disease Control and Prevention LinkPlus probabilistic matching program (CDC 2018). This freely available software is useful for obtaining matches using a probabilistic scoring system developed by Fellegi and Sunter (1969). Important features of LinkPlus include the ability to link by names phonetically (including nicknames), transposition of digits in SSN, and by individual components of date fields (e.g. month of birth).

During this process, it is not unusual to discover errors within or missing key demographic variables, particularly the SSN. Online resources exist which provide qualified entities with access to public records (such as tax assessments, drivers licenses, etc.). The service we utilized most frequently was LexisNexis® Smartlinx® comprehensive person report. By manually searching these records for participants with known errors in identifying data, errors can be corrected and matches can be improved. Vital status may also be obtained this way, as some states provide mortality information through these online resources. However, as this is a time-intensive, manual process, this is only performed after bulk, programmatic tracing options have been exhausted. General

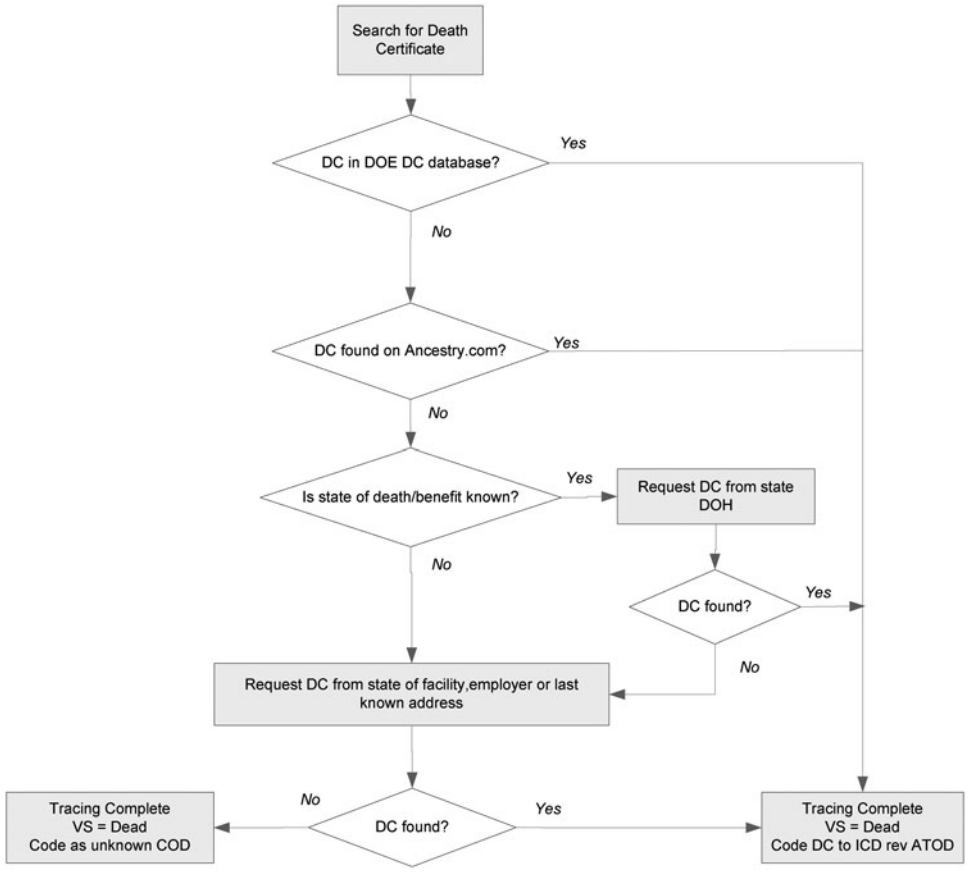

Figure 2. Generalized process to obtain death certificates utilized in the Million Person Study.

web searches for publicly available vital status information, like obituaries, can also be helpful (Sesso et al. 2000).

Corrections or additions to records (either in the participant records or in the IMD record) are entered into dedicated corrections databases. Additional series of matches against the IMD, or submissions to SSA-SER and/or NDI incorporating these corrections are then conducted as needed.

After the matching to the IMD, vital status tracing is considered complete for participants with a matched death date and COD. Remaining participants fall into two groups: (1) Date of death is known but no COD was identified or (2) No vital status determination at all. Those with a date of death available and year of death in 1979 or later are sent to NDI Plus to obtain underlying and contributing causes of death coded to the applicable revision of the ICD. Those with no vital status determination are then sent to SSA-SER to obtain a date of death or verification of alive status, or remain vital status unknown. After the SSA-SER results identify the group of participants with a date of death, those records with death dates 1979 and later are also sent to NDI Plus to obtain causes of death. For those participants verified by SSA-SER as alive after this stage, tracing is complete.

The remaining pool of participants is now those with no match or unknown vital status from the initial linkage to the IMD and the linkage to SSA-SER. This group of workers is sent to NDI Plus for annual searches, beginning with the year last known to be alive or 1979, whichever is later, to obtain both fact and COD.

At this stage, all participants not matched to the IMD, the SSA-SER, or NDI Plus are presumed to be alive. Either a sample or all participants presumed alive (depending on the size of the study population) are individually searched using the LexisNexis® Smartlinx® comprehensive person report. Using alternate methods of searching is often needed when key pieces of identifying information are missing or incorrect. If public records indicate recent activity (e.g. a home sale), and no indication of death is found, it provides additional validation that the participant is alive. If additional data is found (e.g. an alternate name), data fields need corrections, or a date death is provided, these are entered into the corrections databases.

Obtaining cause of death before 1979

Deaths identified prior to 1979 are particularly challenging because the NDI does not include deaths prior to this time. As such, a death certificate must be obtained for all deaths prior to 1979, and the COD must be coded to the ICD revision at the time of death by a trained nosologist. Figure 2 shows the approach MPS investigators used to search for death certificates. The DOE death certificates database was first searched to determine whether a scan of the death certificate is already on hand. Since over 30% of the MPS covers DOE workers, this unique resource is of special value. If a death certificate is not found within the DOE resource, Ancestry.com, which currently has death certificates from 18 states and Puerto Rico, is searched.

Table 1. Summary of vital status tracing activities for selected Million Person Study cohorts.

Study	Total population	Cohort start of follow-up	Vital status year of last follow-up	Death count	Deaths (% of cohort)	COD count	COD (% of deaths)	Confirmed vital status (count)	Confirmed vital status (%)
Atomic veterans (Boice et al.2018, 2019c)	114,270	1945	2010	75,402	66.0	72,803	96.6	110,608	96.8
Nuclear utility workers (Boice et al.2018, 2019a)	145,209	1968/1977	2011	31,952	22.0	31,436	98.4	135,270	93.2
Industrial radiographers (Boice et al.2018, 2019b)	125,867	1968/1977	2011	22,012	17.5	21,598	98.1	110,860	88.1
Mound (Boice et al.2014)	7269	1944	2009	3681	50.6	3629	98.6	6761	93.0
Rocketdyne (Boice et al.2011)	5801	1948	2008	2383	41.1	2372	99.5	5405	93.2
Mallinckrodt chemical works (Golden et al.2019)	2514	1942	2012	1895	75.4	1879	99.2	2493	99.2
Los Alamos National Laboratory (LANL)[a]	23,308	1943	2016+	11,105	47.6[a]	6656	59.9[a]	18,582	79.7[a]

[a]The tracing process for LANL is an example of a work in progress. The values reported here reflect the status to date.

Obtaining approval from SDOH to obtain death certificates can be cost-prohibitive for a number of states, as some states charge based on the number of searches rather than the number of death certificates found and provided. Therefore, a targeted approach was used. The state of death or SSA benefit (if known) is searched first. Next, the state of the facility where the subject was last employed is searched. If a death certificate is obtained, it is then provided to a trained nosologist for coding of the underlying COD and all contributory causes, using the ICD revision in place at the time of death.

Results

Table 1 shows the proportion of subjects successfully traced for vital status in six separate cohorts to date (nuclear power plant workers, industrial radiographers, atomic veterans and workers at Rocketdyne/Atomics International, Mound nuclear facility, and Mallinckrodt chemical works) and one cohort (Los Alamos National Laboratory (LANL workers) in progress. The LANL 'work in progress' indicates the challenges faced to achieve a high location rate after the initial location efforts are made. Within the MPS there are other cohorts in similar stages of tracing, such as the medical radiation workers ($n = 170{,}000$). In the cohorts described in Table 1, the COD ascertainment has been over 96% and verification of vital status has been over 90% of subjects in all but the Industrial Radiographers Cohort, related in part to incomplete identifying information.

Discussion

Other investigators have used a multistage approach for vital status ascertainment in large occupational cohort mortality studies (Buchanich et al. 2005; Hermansen et al. 2009; Wojcik et al. 2010) . We enhanced these approaches in two ways. First, we conducted matches using a collection of state mortality files assembled in-house using both deterministic and probabilistic algorithms. Second, we utilized online service providers such as LexisNexis® Smartlinx® comprehensive person report to access public records to correct and/or confirm key demographic variables used in the linkage process. Using an iterative approach with multiple passes allowed us to incorporate changes and corrections to participant records and to increase the change of obtaining a correct match. In addition, DOE death certificates were extremely valuable to the MPS for obtaining pre-1979 causes of death among facility workers who started working as early as the 1940s.

The completeness of mortality records from the SSA (both in the DMF and SER) varies greatly by age and year of death. Incompleteness may be as high as 40% in recent years (Jones and Vawdrey 2015). However, the SSA data contains the only national source of death information prior to 1979. Therefore, SSA data are very helpful in identifying these early deaths and often provide sufficient information to identify the state of death, from which a death certificate might be requested (Schnorr and Steenland 1997). In the future, investigators may need to rely more heavily on NDI searches, which provide very high-quality matches but can be cost-prohibitive.

While multiple mortality data sources with probabilistic matching algorithms (and manual review) are available to determine fact of death, only one national source (SSA-SER), which uses very strict matching criteria, is currently available to verify alive status. A participant record with common errors such as an alternate, maiden name, or nickname, misspelled last name, a transposed digit within the SSN, or missing month and/or day of birth will likely be probabilistically matched to a SDOH or NDI Plus death record if the participant has died. Because the SSA-SER requires near exact match among SSN, name, and DOB, the participant would not be verified as alive. Therefore, study populations with higher quality or more complete matching variables will have a higher percentage of verified vital status.

The tracing approach presented here for the MPS had the first stage as matching to in-house mortality sources and the second stage as submission to SSA-SER. However, these stages could have been reversed, like Buchanich et al. (2005). Any participants confirmed as alive from SSA-SER could be removed from the tracing process. This approach would reduce the probability of incorrect death matches, and perhaps allow looser matching criteria. Likewise, every mortality resource described herein may not be as useful for cohorts from a different study base. Investigators with occupational cohorts from a single facility might consider obtaining

mortality files from one or surrounding states. The utilization of multiple, overlapping sources of mortality information, extensive data cleaning, and identification of specifically targeted resources able to supplement national databases, should assist any study in obtaining as complete mortality follow-up as possible.

Conclusion

The purpose of this article is to generally describe the comprehensive tracing efforts to locate and obtain COD for one million U.S. radiation workers and veterans, which has been ongoing for decades and involves over 20 different cohorts, some of which requiring specialized approaches. Nonetheless, we hope that some of the approaches described might be applicable to all cohorts, large or small, since the goals of achieving a high follow-up rate are the same, but may require cohort-specific modifications.

Notes on contributors

Michael Mumma is the Director of Information Technology at the International Epidemiology Institute and the International Epidemiology Field Station for Vanderbilt University Medical Center. He has over 20 years of experience in data analysis and conducting epidemiologic investigations.

Sarah Cohen is a Principal Epidemiologist at EpidStat Institute where she directs observational research studies in the areas of pharmacoepidemiology, nutritional epidemiology, and occupational epidemiology as well as leads large data management projects and statistical analyses. She is also an Adjunct Assistant Research Professor of Medicine in the Department of Medicine at Vanderbilt University School of Medicine. She has been a collaborator on the MPS for nearly 20 years, providing analytic support as well as co-authoring numerous publications.

Jennifer L. Sirko is an epidemiologist with over 10 years of experience in occupational and public health research. In addition to her work with the International Epidemiology Institute, she is currently a research officer with the Pew Charitable Trusts.

Elizabeth (Betsy) Dupree Ellis currently works in the Health, Energy and the Environment Program, Oak Ridge Associated Universities. Dr. Ellis does research in Public Health, Occupational Health, and Epidemiology. A major project is contributing to the 'Million Person Study'. She is also active in human subject protection in research.

John Boice is President of the National Council on Radiation Protection and Measurements and Professor of Medicine at Vanderbilt University. He is an international authority on radiation effects and served on the Main Commission of the International Commission on Radiological Protection and on the United Nations Scientific Committee on the Effects of Atomic Radiation. He directs the Million Person Study of Low-Dose Health Effects.

Disclosure statement

The authors report no conflict of interest. The authors alone are responsible for the content and writing of this article.

Funding

This work was supported in part by a grant from the U.S. Nuclear Regulatory Commission [NRC-HQ-60-14-G-0011], a grant from the Centers for Disease Control and Prevention [5UE1EH000989] a grant from the National Aeronautics and Space Administration [Grant No. NNX15AU88G], and grants from the U.S. Department of Energy [Grant No. DE-SC0008944 and Grant No. DE-AU0000042 awarded to the National Council on Radiation Protection and Measurements, which included interagency support from the U.S. Nuclear Regulatory Commission, the U.S. Environmental Protection Agency, and the National Aeronautics and Space Administration). Further, contract support was received by Oak Ridge National Laboratory from the Office of Radiation and Indoor Air, U.S. Environmental Protection Agency, under Interagency Agreement DOE No. 1824 S581-A1, under contract No. DE-AC05-00OR22725 with UT-Battelle; and contract support was received by Oak Ridge Associated Universities from the U.S. Department of Energy under contract No. DE-SC0014664.

ORCID

Elizabeth D. Ellis http://orcid.org/0000-0003-2871-1587

References

Acquavella JF, Donaleski D, Hanis NM. 1986. An analysis of mortality follow-up through the national death index for a cohort of refinery and petrochemical workers. Am J Ind Med. 9:181–187.

Boice JD, Jr, Cohen SS, Mumma MT, Ellis ED, Eckerman KF, Leggett RW, Boecker BB, Brill AB, Henderson BE. 2011. Updated mortality analysis of radiation workers at Rocketdyne (Atomics International), 1948–2008. Radiat Res. 176:244–258.

Boice JD, Jr, Cohen SS, Mumma MT, Ellis ED, Cragle DL, Eckerman KF, Wallace PW, Chadda B, Sonderman JS, Wiggs LD. 2014. Mortality among mound workers exposed to polonium-210 and other sources of radiation, 1944–1979. Radiat Res. 181:208–228.

Boice JD, Jr, Cohen SS, Mumma MT, Hagemeyer D, Chen H, Yoder RC, Dauer LT. 2019a. Leukemia among early nuclear power plant workers employed between 1957 and 1984 in the United States. Int J Radiat Biol. (Submitted).

Boice JD, Jr, Cohen SS, Mumma MT, Hagemeyer D. 2019b. Leukemia among industrial radiographers employed between 1940 and 2010 in the United States. Int J Radiat Biol. (this issue).

Boice JD, Jr, Ellis ED, Golden AP, Girardi DJ, Cohen SS, Chen H, Mumma MT, Shore RE, Leggett RW. 2018. The past informs the future: an overview of the million worker study and the Mallinckrodt chemical works cohort. Health Phys. 114:381–385.

Boice, JD, Jr, Cohen SS, Mumma MT, Chen H, Golden AP, Beck HL, Till JE. 2019c. Mortality among US military participants at seven above-ground nuclear weapons test series. Int J Radiat Biol. (this issue).

Boyko EJ, Koepsell TD, Gaziano JM, Horner RD, Feussner JR. 2000. U.S. department of veterans affairs medical care system as a resource to epidemiologists. Am J Epidemiol. 151:307–314.

Boyle CA, Decouflé P. 1990. National sources of vital status information: extent of coverage and possible selectivity in reporting. Am J Epidemiol. 131:160–168.

Buchanich JM, Dolan DG, Marsh GM, Madrigano J. 2005. Underascertainment of deaths using social security records: a recommended solution to a little-known problem. Am J Epidemiol. 162: 193–194.

Campbell KM, Deck D, Krupski A. 2008. Record linkage software in the public domain: a comparison of link plus, the link king, and a 'basic' deterministic algorithm. Health Informatics J. 14:5–15.

[CDC] Centers for Disease Control and Prevention. 2018. National program of cancer registries (NPCR). Registry plus™ link plus. National death index. Hyattsville (MD): CDC. (http://www.cdc.gov/cancer/npcr/tools/registryplus/lp.htm)

Cowper DC, Kubal JD, Maynard C, Hynes DM. 2002. A primer and comparative review of major US mortality databases. Ann Epidemiol. 12: 462–468.

Curb JD, Ford CE, Pressel S, Palmer M, Babcock C, Hawkins CM. 1985. Ascertainment of vital status through the national death index and the social security administration. Am J Epidemiol. 121:754–766.

Fellegi IP, Sunter AB. 1969. A theory for record linkage. J Am Stat Assoc. 64:1183–1210.

Fillenbaum GG, Burchett BM, Blazer DG. 2009. Identifying a national death index match. Am J Epidemiol. 170:515–518.

Golden AP, Ellis ED, Cohen SS, Mumma MT, Leggett RW, Wallace PW, Girardi D, Watkins JP, Shore R, Boice JD. Jr. 2019. Updated mortality analysis of the Mallinckrodt uranium processing workers, 1942–2012. Int J Radiat Biol. (this issue).

Hermansen SW, Leitzmann MF, Schatzkin A. 2009. The impact on national death index ascertainment of limiting submission to social security administration death master file matches in epidemiologic studies of mortality. Am J Epidemiol. 169:901–908.

Jones B, Vawdrey D. 2015. Measuring mortality information in clinical data warehouses. AMIA Jt Summits Transl Sci Proc. 2015:450–455.

LexisNexis®. 2018. LexisNexis public records collection, available via Smartlinx® comprehensive person report.

Maynard C, Chapko MK. 2004. Data resources in the department of veterans affairs. Diabetes Care. 27:B22–B26.

National Center for Health Statistics. 2018. Chapter 4, Assessing NDI Output. http://www.cdc.gov/nchs/data/ndi/ndi_users_guide_chapter4.pdf.

Page WF, Mahan CM, Kang HK. 1996. Vital status ascertainment through the files of the department of veterans affairs and the social security administration. Ann Epidemiol. 6:102–109.

Patterson BH, Bilgrad R. 1986. Use of the national death index in cancer studies. J Natl Cancer Inst. 77:877–881.

Schall LC, Buchanich JM, Marsh GM, Bittner GM. 2001. Utilizing multiple vital status tracing services optimizes mortality follow-up in large cohort studies. Ann Epidemiol. 11:292–296.

Schnorr TM, Steenland K. 1997. Identifying deaths before 1979 using the social security administration death master file. Epidemiology. 8: 321–323.

Sesso HD, Paffenbarger RS, Lee IM. 2000. Comparison of national death index and world wide web death searches. Am J Epidemiol. 152: 107–111.

Social Security Administration. 2018. Service to epidemiological researchers to provide vital status data on subjects of health research. https://www.ssa.gov/policy/about/epidemiology.html.

Wojcik NC, Huebner WW, Jorgensen G. 2010. Strategies for using the national death index and the social security administration for death ascertainment in large occupational cohort mortality studies. Am J Epidemiol. 172:469–477.