Policy Commentary

Consequences for Healthcare Quality and Research of the Exclusion of Records From the Death Master File

Briget da Graca, JD, MS; Giovanni Filardo, PhD, MPH; David Nicewander, MS

In November 2011, the Social Security Administration removed ≈5% of death records from its Death Master File and started excluding ≈40% of new death records, having determined that data submitted electronically by states cannot be publicly shared. Before this determination, the Death Master File provided an accessible source of national vital status data with a short time lag and high specificity and sensitivity and was routinely used by healthcare researchers and hospitals to determine study participants’ survival and to monitor postdischarge outcomes. Its effective loss means comparative effectiveness studies will be unnecessarily delayed, more costly, or unfeasible. Likewise, timely identification and correction of poor hospital performance will be more difficult, undermining the safety and quality of care and threatening hospital financing as the Centers for Medicare and Medicaid launch the Readmissions Reduction Program in October 2012 and link reimbursement to 30-day mortality under the Value-Based Purchasing Program in 2013. In summary, the action of the Social Security Administration will substantially hamper healthcare research and quality. We describe the origins of the Death Master File and the basis for excluding electronically submitted state data. We then examine the consequences for healthcare research and operations, consider alternative sources, and evaluate possible mechanisms to restore a timely national data source.

On November 1, 2011, the Social Security Administration (SSA) removed ≈5% of the data in its publicly available Death Master File (DMF) and stopped reporting ≈40% of new deaths.1 The SSA explained that it had determined that §205(r) of the Social Security Act (added by the Act of April 20, 1983)2 prohibits the disclosure of state records that have determined that data submitted electronically by states cannot be publicly shared. Before this determination, the Death Master File provided an accessible source of national vital status data with a short time lag and high specificity and sensitivity and was routinely used by healthcare researchers and hospitals to determine study participants’ survival and to monitor postdischarge outcomes. Its effective loss means comparative effectiveness studies will be unnecessarily delayed, more costly, or unfeasible. Likewise, timely identification and correction of poor hospital performance will be more difficult, undermining the safety and quality of care and threatening hospital financing as the Centers for Medicare and Medicaid launch the Readmissions Reduction Program in October 2012 and link reimbursement to 30-day mortality under the Value-Based Purchasing Program in 2013. In summary, the action of the Social Security Administration will substantially hamper healthcare research and quality. We describe the origins of the Death Master File and the basis for excluding electronically submitted state data. We then examine the consequences for healthcare research and operations, consider alternative sources, and evaluate possible mechanisms to restore a timely national data source.

What Is the DMF?
The SSA DMF is a public-use file that the SSA created to meet Freedom of Information Act requests from members of the public seeking vital status data.6-7 The DMF contains the Social Security number; first, middle, and last names; dates of birth and death; and last known ZIP code for individuals who the data of SSA show are deceased.8,9

The SSA receives death data from a wide variety of sources but primarily from family members of the deceased, funeral homes, financial institutions, and the states.9 On receipt, these deaths are annotated in the numerical identification (NUDIMENT) file.10 The SSA DMF is a public-use file that the SSA created to meet Freedom of Information Act requests from members of the public seeking vital status data.6-7 The DMF contains the Social Security number; first, middle, and last names; dates of birth and death; and last known ZIP code for individuals who the data of SSA show are deceased.8,9

As required by law, the SSA shares the full set of death data with federal benefit-paying agencies.10 The publicly available set—the DMF—is distributed through the National Technical Information Service of the Department of Commerce.5 National Technical Information Service sells the DMF to >450 entities categorized as (1) federal, state, and local governments; (2) industry (including financial, investigative, credit reporting, and medical research or healthcare organizations); and (3) public customers, including private individuals and genealogical services.6,8 In 2011, the SSA concluded that it had been impermissibly including state-owned data in the DMF since 2002, prompting the removal of ≈4.2 million records from the publicly available DMF and the exclusion of ≈40% of new death reports being submitted to the SSA.1

The Basis for the SSA’s 2011 Exclusion of Data From the DMF
The US federal government is one of limited powers, with all powers not specifically assigned to it reserved to the states11;
Consequences of the Reduced DMF Content for Healthcare Research

The DMF has been a valuable tool for healthcare research examining mortality or survival because it captures deaths across all US states and territories, is available at low or no cost, and has relatively high sensitivity (≥86% overall but >95% in older generations), high specificity (100%), and only a 4 to 6-month time lag in death reporting. All alternative sources are less timely, less complete, and more expensive. Critical epidemiologic, comparative effectiveness, and clinical trials research, as well as surveillance programs, depend heavily on timely and accurate assessments of vital status, and the limited funding available for such research requires the assessment not to consume inordinate amounts of the research budget or time. Particularly, when looking at outcomes years after the initial treatment, a resource like the DMF is necessary because tracking study subjects who have moved and changed their names would be costly, time-consuming, and not necessarily successful.

A timely vital status assessment is critical to provide patients with relevant data to inform their healthcare decisions. When availability of vital status data is delayed (such as the 2-year delay reported for the National Death Index [NDI]), major issues follow. First, it is impossible to timely examine the effectiveness of available treatments or technologies, which is essential because efficacy demonstrated in clinical trials does not always translate into real-world outcomes. Second, when short-term outcomes that can be assessed through direct observation (eg, operative mortality) show greater benefit with 1 treatment but longer-term outcomes requiring postdischarge vital status assessment (eg, 1-year mortality) favor another, the delayed assessment of the longer-term outcomes leads to unnecessary loss of life. The recent American College of Cardiology Foundation-The Society of Thoracic Surgeons Collaboration on the Comparative Effectiveness of Revascularization Strategies (ASCERT) study provides an example of when this might arise: Although perioperative risks for elderly patients undergoing cardiac revascularization are lower for percutaneous coronary intervention, coronary artery bypass graft surgery offers significantly better survival at ≥2 years.

The demise of the DMF is ironic in that it follows on the heels of substantial federal funding for comparative effectiveness research: The American Recovery and Reinvestment Act of 2009 appropriated $1.1 billion to the Department of Health and Human Services for comparative effectiveness research, and the Patient Protection and Affordable Care Act of 2010 established the Patient-Centered Outcomes Research Institute to focus on comparative clinical effectiveness research, supported by a trust fund to which $50 million was appropriated for fiscal year 2011 and $150 million for fiscal year 2012. Surveillance programs—critical for understanding disease progression and patients’ outcomes so that a coordinated system of care that maximizes benefit across the population can be developed—also depend on the availability of timely vital status data. Examples of such programs for which researchers already report feeling the loss of the DMF include the Nurses’ Health Study (the Harvard-based study of cancer prevalence among >200 000 women that has been ongoing for 36 years and formed the basis for ≥50 publications in the peer-reviewed literature) and the Framingham Heart Study (the Boston University–based study that has been ongoing since 1948 and from which several hundred peer-reviewed publications have originated). The senior project manager of the Nurses’ Health Study has been quoted as saying the new policy of the SSA had “thrown us back to the pre-Internet era, where you’d start looking in the phone book for someone with a similar name and sending out a bunch of letters,” and the director of operations for the Framingham Heart Study has described it as the loss of a very valuable tool that is wasting research dollars. As the examples of these renowned studies demonstrate, with the exclusion of EDR-submitted death reports from the DMF, critical healthcare studies will be more difficult, expensive, time-consuming, and delayed. This will
reduce the contemporary evidence available to patients and clinicians facing treatment decisions, and to policymakers facing decisions about the structure and coverage of Medicare and similar programs and the operation of the healthcare system. Furthermore, it raises a substantial barrier to the United States remaining a leader in comparative effectiveness research.

Consequences of the Reduced Content of the DMF for Hospitals

Hospitals subscribe to the DMF to monitor postdischarge patient outcomes to assess and improve the quality, safety, and efficiency of care. Key outcome measures for conditions with high short-term mortality depend on an accurate assessment of postdischarge vital status—specifically, 30-day mortality and 30-day readmission rates (the latter can be biased if the competing risk of death is not accounted for). Hospitals match patient identifiers with the DMF to estimate and monitor 30-day mortality and readmission rates. These rates are used by hospitals’ patient safety and quality departments to identify areas of low performances, which can then be targeted for improvement through the implementation of existing safety and quality programs and the development of new ones. The 4- to 6-month lag in reporting death by the DMF facilitated relatively timely identification and correction of poor performance, minimizing patient risk and harm. But with the exclusion of EDR-submitted data from the DMF, hospitals must rely on alternative sources, which have longer time lags. This delays identification of safety and quality issues, exposing more patients to risk in the interim and allowing problems to become more deeply entrenched and harder to correct.

In-hospital mortality provides an incomplete picture of the quality and safety of hospital care and an inaccurate estimate of a hospital’s publicly reported 30-day postdischarge mortality. This could have substantial financial consequences for hospitals with the advent of the Centers for Medicare and Medicaid Services Readmissions Reduction Program in October 2012 and the use of 30-day postdischarge mortality as one of the measures determining hospital reimbursement. The Centers for Medicare and Medicaid Services Value-Based Purchasing Program starting in 2013. Although the access of Centers for Medicare and Medicaid Services to mortality data needed to determine the performance of the hospitals has not changed, the hospitals now lack timely data to monitor 30-day mortality (and to take the competing risk of death into account in monitoring readmission rates) and to estimate the corresponding reimbursement they will receive. The inability to monitor postdischarge mortality also prevents early identification of poor performance, which would enable rapid responses to minimize the period of unnecessary patient risk and reduced reimbursement. Although only 1% of a hospital’s base operating diagnosis-related group payments are initially at risk under each of these programs, the stakes will rise in future years, with the planned progression placing up to 3% of Medicare reimbursements at risk based on readmissions performance by October 2014 and a further 2% based on quality of care (including mortality) by 2016. These percentages can nonetheless amount to substantial sums. For example, the Wexner Medical Center of The Ohio State University estimates that it will lose $700,000 next year because of the 0.64% penalty assessed under the Hospital Readmissions Reduction Program in October 2012.

Alternative Sources of Mortality Data

Only states have the authority to collect death data. State vital statistics offices are the first point of collection for death data and are an alternative to the DMF. But to assess vital status nationally, a separate request must be submitted to each of the 57 vital records jurisdictions. Each jurisdiction cross-references the personal health information provided with its death registry to determine whether any subjects died there—a process taking up to 12 months. Furthermore, restrictions on data release vary between states, making it uncertain that data can be obtained for all the purposes for which the DMF was used. Insofar as mortality data needs of researchers and hospitals can be met with in-state death records, this may give those operating in states with less restrictive policies on the release of vital status records an advantage in the wake of the loss of the DMF. However, other factors such as how the data are collected and stored would also play into any relative advantage 1 state may have over another with respect to being able to provide timely death data.

The NDI, available through the Centers for Disease Control and Prevention, is based on monetary contracts with the states to provide their death registries annually. But the NDI is “available to investigators solely for statistical purposes in medical and health research” and “not accessible to organizations or the general public for legal, administrative, or genealogy purposes.” Thus, the NDI may be a viable substitute for the DMF for comparative effectiveness research and surveillance programs but appears not to be available to healthcare organizations monitoring postdischarge mortality and readmission rates. State data are harvested 12 months after the end of each calendar year, and the NDI needs ≈1 year to organize, verify, audit, and enter the data, resulting in a 2-year time lag. Accessing the NDI requires an application describing how the information will be used, which requires an additional 2 months for review. If approved, the NDI provides vital status assessment data for individuals whom the investigator listed, charging a service fee plus a fee per record for each year searched. For large studies and long-term follow-up, these costs accumulate rapidly. Although the DMF also carries a subscription fee, it allows researchers or organizations to run their own queries, with no need for external requests or transfers of patient data. Therefore, despite the greater accuracy of the NDI, the costs, time lag, data transfer risks, and exclusive use for research make it an unsuitable substitute for many of the uses to which the DMF has been applied.

A third means of obtaining mortality data is through the Vital Status Service of the SSA; however, like the DMF, it excludes data obtained from states under §205(r) of the Social Security Act.

Finally, researchers and healthcare organizations interested only in the limited population of Medicare beneficiaries can purchase vital status data stored in the Medicare Master Beneficiary Summary File or Vital Status File. The former is split into calendar year files and has a 9-month lag from the close of the most recent calendar year; the latter is a current file, unrestricted by calendar year and containing the latest
available information.\textsuperscript{32} However, because these data include beneficiary identifiers, they are subject to the Privacy Act of 1974, Health Information Portability and Accountability Act, and other federal government rules and regulations and can be released only after determination that the criteria for release have been met and the proposed use falls under the permitted uses and disclosures and after finalization of a data use agreement.\textsuperscript{33} Because each data use agreement is study specific, a new application (detailing the research purposes and methods, data management and privacy safeguards, and funding source) is required not only for each study requiring new data but also for each reuse of previously purchased data for a purpose not specified in the existing agreement.\textsuperscript{34} Costs vary according to files requested, the nature and size of the cohort requested (standard 5\% or 20\% random sample versus custom cohort or 100\% sample), and the years of data requested.\textsuperscript{15} Alternatively, for hospitals and healthcare organizations seeking data only for performance measurement purposes, and that are able to demonstrate existing expertise in performance measurement, the ability to combine Medicare data with existing claims data, a process allowing providers to review and correct their performance reports, and adherence to rigorous data privacy and security procedures, standardized extracts of Medicare claims data are available under §10332 of the Patient Protection and Affordable Care Act of 2010; 30-day mortality and readmission rates are included in the standard measures developed for this Qualified Entity Program.\textsuperscript{36}

**What Are Our Options?**

One substitute has already been suggested: that the Centers for Disease Control and Prevention distribute mortality data to researchers monthly by subscription, containing the NDI data known to date and supplemented by up-to-date SSA data.\textsuperscript{3} This would address the problem of the historical data removed from the DMF but offers no substitute for DMF data that originated from state reports for the most recent 2 years or for the \( \approx 40\% \) of new deaths excluded as state-owned data. Moreover, because states submit data to the NDI only once annually,\textsuperscript{18} the monthly updates would not offer a more complete database than a search of the NDI followed by a search of the (incomplete) DMF for the 2 years not yet covered—unless the state contracts and process by which the NDI is updated were revised for closer to real-time reporting. On a more positive note, the Centers for Disease Control and Prevention is working with states offices to adopt electronic filing systems that can keep the NDI more current, but according to its director of vital statistics, it will not be able to reduce costs.\textsuperscript{21-37} However, although this solution might mitigate the loss of EDR-submitted data from the DMF for researchers, the restriction on the use of NDI data to research appears to exclude use by healthcare organizations monitoring quality of care.

It may, therefore, indeed require an act of Congress, powered by “a grass-roots uprising of the biomedical community, medical societies, lobbyists, and congress-persons”\textsuperscript{3} to restore the DMF as a resource for healthcare research and quality monitoring.\textsuperscript{3} However, the hope for congressional rescue faces a substantial barrier in the introduction of 2 bills in late 2011 motivated by the use of DMF data in tax fraud,\textsuperscript{38} which would further reduce public access: H.R. 3475 would exempt the entire DMF from the Freedom of Information Act,\textsuperscript{39} and S. 1534 would exempt all data for individuals who died during the year the request was made for that year and the succeeding year, allowing only limited exceptions for fraud prevention.\textsuperscript{40} Although the extent to which public access to DMF data creates a risk of tax fraud is beyond the scope of this article, that risk may be sufficient to justify removing the DMF from general public access. If that were found to be the case, a better alternative might be along the lines of that advocated by the Society of Thoracic Surgeons: not opposing the general removal of the DMF from public access by H.R. 3475 but expanding the permitted uses of the state-provided death records in §205(r) of the Social Security Act to include other statistical and research activities conducted by medical, scientific, or public health researchers in accordance with the Federal Common Rule (46 C.F.R. §46.101) and the applicable privacy and security rules issued under the Health Information Portability and Accountability Act of 1996 (45 C.F.R. Pt. 164 and § 512(ii)), and subject to review and approval by an institutional review board registered with the Department of Health and Human Services Office of Human Subjects Protection and the Food and Drug Administration and accredited by the Association for the Accreditation of Human Subjects Research Protection Programs.\textsuperscript{41}

Such an approach appears to strike a balance between the privacy concerns associated with tax fraud and the need for an inexpensive timely source of mortality data for healthcare researchers and organizations. Moreover, in substance, it appears to differ from the agreement under which state death data are released in the NDI only in that the states would not be paid for submitting the data to the SSA. Although half the states do not give the SSA permission for public dissemination of the death data they submit,\textsuperscript{18} all 50 states allow data submitted to the NDI to be used for research purposes. We therefore urge healthcare organizations and researchers not only to lend their support to the federal agencies with stakes in medical research that are reportedly “lobbying Social Security officials to consider a compromise,”\textsuperscript{21} but also to lend their weight to congressional action to restore and protect this important data source for uses that will affect the quality of care we are able to provide and receive now and in years to come.

**Sources of Funding**

This work was supported in part by the Bradley Family Endowment to Baylor University Medical Center.

**Disclosures**

None.

**References**


11. U.S. Const. amend. X.
12. Hearing Before the Subcommittee on Social Security of the House Committee on Ways and Means, 112th Cong. Statement for the record by Patricia W. Potrzebowski, PhD, on behalf of the National Association for Public Health Statistics and Information Systems. 2012.
34. Mann E. Tips for Completing a Research Identifiable Data Request Pack-
et. 2011; http://www.resdac.org/resconnect/articles/153. Accessed Octo-

Key Words: comparative effectiveness ■ death master file ■ hospital performance ■ mortality ■ safety ■ surveillance
Consequences for Healthcare Quality and Research of the Exclusion of Records From the Death Master File
Briget da Graca, Giovanni Filardo and David Nicewander

doi: 10.1161/CIRCOUTCOMES.112.968826
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/6/1/124

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org//subscriptions/