ASPE
ASSISTANT SECRETARY FOR PLANNING AND EVALUATION

OS PCOR TF
OFFICE OF THE SECRETARY
PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND

REPORT

Mortality Data Linkages for Research

Prepared for
the Office of the Assistant Secretary for Planning and Evaluation (ASPE)
at the U.S. Department of Health & Human Services

September 2024

Office Of The Assistant Secretary For Planning And Evaluation

The Assistant Secretary for Planning and Evaluation (ASPE) advises the Secretary of the U.S. Department of Health and Human Services (HHS) on policy development in health, disability, human services, data, and science; and provides advice and analysis on economic policy. ASPE leads special initiatives; coordinates the Department's evaluation, research, and demonstration activities; and manages cross-Department planning activities, such as strategic planning, legislative planning, and review of regulations. Integral to this role, ASPE conducts research and evaluation studies; develops policy analyses; and estimates the cost and benefits of policy alternatives under consideration by the Department or Congress.

The Office Of Health Policy

The Office of Health Policy (HP) provides a cross-cutting policy perspective that bridges Departmental programs, public and private sector activities, and the research community, in order to develop, analyze, coordinate, and provide leadership on health policy issues for the Secretary. HP carries out this mission by conducting policy, economic, and budget analyses; assisting in the development and review of regulations; assisting in the development and formulation of budgets and legislation; and assisting in survey design efforts, as well as conducting and coordinating research, evaluation, and information dissemination on issues relating to health policy.

Office Of The Secretary – Patient-Centered Outcomes Research Trust Fund

The Office of the Secretary – Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) was established as part of the 2010 Patient Protection and Affordable Care Act and is charged to build data capacity for patient-centered outcomes research. Coordinated by ASPE on behalf of the Department, OS-PCORTF has funded a rich portfolio of projects to meet emerging HHS policy priorities and fill gaps in data infrastructure to enhance capabilities to collect, link, and analyze data for patient-centered research. For more information visit https://aspe.hhs.gov/collaborations-committees-advisory-groups/os-pcortf

This report was funded by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) under Contract Number HHSP233201500038I of the HHS Office of the Assistant Secretary for Planning and Evaluation (ASPE). The work was carried out by RAND with contributing authors as shown below. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of HHS, ASPE, RAND, or the other authors' institutions. Readers should not interpret any statement in this product as an official position of ASPE or of HHS.

Contributing Authors

RAND Health Care: Sara Heins, Rachel Burns, Joachim Hero, Baqir Fateh, Stephanie Walsh
Johns Hopkins Bloomberg School of Public Health: Alexander McCourt
ASPE: Joel Dubenitz
National Institutes of Health (NIH): Michael Schoenbaum

Suggested Citation: Heins, S., Burns, R., Hero, J., Fateh, B., Walsh, S., McCourt, A., Dubenitz, J., Schoenbaum, M. Mortality Data Linkages for Research. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, September 2024.

Acknowledgments

We are grateful to the technical expert panelists for their input and to Joshua Breslau, Christine Eibner, and Paul Koegel (RAND); David Swedler (University of Massachusetts); and Sara Couture (ASPE) for their reviews. We thank Leslie Canterbury, Stephanie Lonsinger, and Lisa Turner for administrative and editorial assistance.

EXECUTIVE SUMMARY

Background and Approach

Understanding risk factors for death is a key aspect of epidemiological research and public health practice, and linkable mortality data are a necessary component of survival research. Mortality data for the United States are owned by 57 separate jurisdictions (50 states, five territories, the District of Columbia, and New York City) and compiled by the National Center for Health Statistics as national de-identified statistical files into a national restricted use, linkable mortality dataset, the National Death Index (NDI). There are several documented barriers to obtaining linkable mortality data for survival research, including the lack of a centralized source of information about individual state policies and procedures governing data access and use. This research attempts to fill that knowledge gap via a database of publicly available information from state legislatures, jurisdictions' health department or vital statistics websites, and other sources of information.

The study team conducted a systematic search and review of formal policies pertaining to the access and use of mortality data (i.e., statutes and regulations), as well as policies, processes, and application materials from the responsible agency or department in each of the 57 mortality-reporting jurisdictions. The team then developed a list of policy dimensions through an initial scan of a sample of jurisdictions' vital statistic webpages and applications for access to mortality data, a review of the NDI data application, and consultation with the Assistant Secretary for Planning and Evaluation, a technical expert panel (TEP), and a public health lawyer. The team abstracted and reviewed descriptions of available data; information about the application and review process; legal, ethical, and privacy review requirements; permitted and disallowed uses of the data; data protection requirements; and data linkage policies. The team compiled this information into a database and reviewed it for accuracy and completeness. Finally, the team sent the abstracted jurisdiction-level information to a vital statistics representative, identified by a member of the TEP, from each jurisdiction to ask for additional information, sources, or corrections. The team incorporated relevant feedback into the database.

This report summarizes the methodology used, the sources and content of the data that were collected, and implications for researchers, including use cases and considerations for using mortality data for patient-centered outcomes research (i.e., research that helps patients and their caregivers make informed health care decisions).

Results

Search Results

The amount of information on policies varied considerably by jurisdiction (e.g., we were able to abstract information for more than 75 percent of the data fields for three jurisdictions and less than 10 percent of the data fields for three other jurisdictions). There was also notable variation in the amount of information that we were able to find for each data element (e.g., we were able to identify a description of permitted uses of mortality data in all but four jurisdictions, but very few jurisdictions specified any information on restrictions on which causes of death could be released or whether certain users were disallowed). Below we summarize our findings on the six primary policy dimensions that we identified.

Description of Available Data

Most jurisdictions indicated that they provided standard death certificate data elements,[1] but several did not release Social Security numbers (SSNs) or other identifiers to researchers. Most jurisdictions did not specify the available years of data or only indicated the length of time that must elapse before death records become publicly available, which ranged from 25 to 75 years. Some jurisdictions specified that preliminary data were available (sometimes referred to as "provisional"), but most jurisdictions did not specify the availability of preliminary data. When jurisdictions described data gaps or completeness, they included descriptions of when death certificate forms were changed or provided a general statement noting that available data elements vary across years.

Application and Review Process

The study team identified some level of detail on the application and review process for acquiring mortality data in all but five jurisdictions. However, some of this information was only available through requesting application materials from the jurisdiction. Applications commonly require researcher information and credentials, a description of the study and associated need for the data, specification of requested variables and years of data, information on prior or current Institutional Review Board (IRB) activity, and data use agreements (DUAs). Jurisdiction resources commonly noted that clarification or additional information may be requested following the initial review of application materials.

Most jurisdictions provided instructions on the elements to include in applications. Jurisdictions that provided a timeline for the application process gave a wide range of expected processing times (from weeks to several months), commonly noting that time frames depend on the nature and complexity of the request. There was also variation in fee structures for use of mortality data, including fees per dataset, fees per record, fees per hour of data preparation and or analysis, annual ongoing charges, and flat fees. Some jurisdictions did not charge fees. Most jurisdictions specified that digital data transfers (using Secure File Transfer Protocol [SFTP]) were available.

Legal, Ethical, and Privacy Review Requirements

Most jurisdictions specified that IRB review is required to access mortality data, whether from the applicant's own institution, a jurisdictional IRB, or both. Among most states in which the IRB review requirement was not specified, several required a review from other bodies, such as a data council, a legal reviewer, or a state registrar. In some jurisdictions, additional review could be required by a privacy board or its equivalent.

Permitted and Disallowed Use and Users

No states specified a maximum number of records that they would release, but a few states noted that they would only release the minimum number of records necessary to answer the proposed research question. No jurisdictions put restrictions on the release of specific causes of death (e.g., drug overdose, suicide). Other restrictions included time restrictions on access to the data, not releasing individual-level data if aggregated data could meet research needs, and only sharing record-level information with the consent of the immediate family of the deceased.

Nearly all jurisdictions explicitly allowed mortality data to be used for research. The most frequently cited disallowed use was for commercial purposes.

Most jurisdictions specified allowed users, and more than half of these specified government agencies and researchers. Other specified allowed users include public or private agencies, government contractors, individuals, and health care providers. However, it was often not clear whether all users could access individual records and all available variables. Four jurisdictions specified that commercial firms or agencies were not

allowed to access mortality data.

Data Protection Requirements

Several jurisdictions specified restrictions on the dissemination of research results, including required approval of dissemination products by the department or agency, notification of publication, or acknowledgement of the department or agency as the source of data. Several jurisdictions specified that releasing personally identifiable information or small cell counts is prohibited.

Most jurisdictions did not specify whether open-ended access to data was allowed, but those who did specify did not permit it unless a new or renewal application was submitted and approved. Most jurisdictions also did not specify whether follow-back investigations were permitted.

Most jurisdictions that required a DUA or other attestation to obtain data also required that applicants submit a data safeguarding plan to explain how they would ensure the security of the provided data. While most jurisdictions had requirements in the DUA or other attestation that data be kept secure, most jurisdictions did not specify where data should be stored or accessed. Most jurisdictions required a data disposition plan.

Data Linkage Policies

Among jurisdictions explicitly allowing data linkage, there was considerable variation in available linkage methodologies. For all jurisdictions allowing researchers to link data themselves, researchers needed to complete some type of application and/or DUA, typically outlining the linkage variables, the steps for data linkage, and the intended uses of the linked data. Most jurisdictions did not specify the variables or methods that could be used for linkage.

No jurisdictions explicitly disallowed all types of data linkage. Those that disallowed researchers to link data themselves could accommodate requests for the department to link the data and return de-identified data to the researchers. Some indicated that researchers could request identifying information that could potentially be used for data linkage. Several jurisdictions' data applications have questions about data linkage, but it was unclear whether researchers could perform the data linkages themselves based on the application alone.

A few states routinely link mortality data with other datasets and, under some circumstances, make these linked data available to researchers.

Implications for Researchers

After reviewing available information across the six policy dimensions described above, for each policy dimension, we identified those features that most readily facilitate survival research: (1) standard death certificate elements (other than SSNs) and data from the past year available to researchers; (2) a transparent application and review process; (3) transparent legal, ethical, and privacy review requirements; (4) explicit permission for data to be used for researcher-directed research (as opposed to jurisdiction-solicited research); (5) transparent data protection policies; and (6) explicit permission for individual-level data linkage. We identified five model jurisdictions (KY, NC, OR, PA, and WA) that have all six characteristics. While each of these jurisdictions has policies facilitating the use of mortality data for PCOR, researchers may still be successful in using mortality data from other jurisdictions depending on the specific research questions and needs.

Conclusion

Nearly all jurisdictions explicitly allowed mortality data to be used for research, but most lacked detailed, publicly available information on the processes for accessing and linking these data. We identified five model

jurisdictions with policies allowing for the use of PCOR data. This report shows that research using linkable mortality data is feasible, but improved transparency in data access policies would increase the volume and diversity of research using these data.

TABLE OF CONTENTS

TABLE OF EXHIBITS

CHAPTER 1. INTRODUCTION

Background

Survival is a key epidemiological outcome, and access to mortality data is essential for studying risk factors for death due to a myriad of causes, from coronavirus disease 2019 (COVID-19) to suicide and overdose. Mortality data are also critical to health equity research. For example, linked vital record data were used to understand and address the root causes of racial and ethnic disparities in mortality early in the COVID-19 public health emergency.[2] However, there are several documented barriers to conducting survival research, including challenges with obtaining linkable mortality data.

The National Death Index (NDI) is the only comprehensive, linkable source of national mortality data available to approved researchers for approved projects. These projects include patient-centered outcomes research (PCOR), which is defined as "research that helps people and their caregivers communicate and make informed health care decisions."[3] Linking mortality data with other data sources measuring health outcomes supports research that provides patients and their caregivers with more complete information to make such decisions. Examples of projects using NDI data include long-term follow-up after cancer clinical trials,[4] tracking long-term mortality among workers in industries with high radiation exposure,[5] and examining factors associated with recurring firearm injuries and death.[6] The National Center for Health Statistics (NCHS) partners with state vital record offices to systematically collect information on the state in which a death occurred, the date of death, death certificate numbers, and cause of death; these data are compiled into the NDI.[7] Although the NDI was created with the intention of having a centralized source of linkable mortality data at the national level, researchers have reported that these data are costly to obtain ($350 for a base service charge, plus $0.15 to $5.00 per record, depending on the type of search[8]) and that the application process is burdensome, confusing, and time-consuming.[9] The application must be completed by the principal investigator,[10] and the approval process and time between approval and receipt of data are extensive (the review process alone takes typically two to three months[11]). Researchers cannot link the data themselves; instead, they must submit a dataset with subject identifiers (e.g., first and last name and Social Security number [SSN], first and last name and month and year of birth, or SSN and full date of birth and sex) with a transmittal form to receive encrypted files with information about linkage and matching from NDI staff.

Criticism among epidemiologists about bureaucratic barriers to research is not new; nor is the discussion of balancing the public health gains of population health research with the legitimate concerns about privacy protections and risks of unauthorized disclosures.[12-14] The Common Rule does not apply to research involving only decedents; however, NDI data are subject to other statutory protections, including the Privacy Rule and Section 308(d) of the Public Health Service Act.[15] NCHS employees are also subject to the Confidential Information Protection and Statistical Efficiency Act or CIPSEA,[16] which requires them to take an oath and face jail terms of up to five years and/or a fine of up to $250,000 in the event of any willful disclosure of identifiable information.[17] These laws are intended to protect deceased individuals' data from unauthorized uses, identity theft, and/or fraud; however, concerns remain that existing data restrictions prevent and/or inhibit legitimate research activities, even when appropriate data protection measures are in place.[14]

As an alternative to using the NDI, researchers may go directly to jurisdictions for linkable mortality data. Linkable mortality data in the United States are owned by 57 separate jurisdictions: 50 states (denoted by the U.S. Postal Service's two-letter abbreviations), Puerto Rico (PR), Guam (GU), Virgin Islands (VI), American Samoa (AS), Northern Mariana Islands (MP), the District of Columbia (DC), and New York City (NYC). Each jurisdiction has its own policies on accessing and using vital record data for research), and these individual policies inform the policies for the NCHS's NDI, because the national datasets must adhere to the most restrictive of the policies of the included jurisdictions. Going directly to jurisdictions may be less restrictive and sometimes allows researchers to link data themselves. However, because each jurisdiction has a separate

process for accessing data, approaching jurisdictions directly may only be practical if data from only one or two jurisdictions are required. Examples of studies using state mortality data linked with other data sources include linkage with hospital discharge data to study long-term mortality outcomes following hospitalization for strokes in Massachusetts,[18] linkage with cancer and transplant registries to assess long-term mortality outcomes in California,[19] and linkage with immunization information systems to more accurately determine population-level immunization rates in Iowa.[20] However, details about the mortality data use policies of these jurisdictions are not readily available in a central location, and there is no standardization across jurisdictions in how information about data access is presented to potential users. This lack of centralized policy information may be why research studies using jurisdictions' linked vital records commonly include a government partner from the jurisdiction.[21]

NCHS previously partnered with the National Association of Public Health Statistics and Information Systems (NAPHSIS) to conduct a survey of members about their perceptions of barriers to using linkable mortality data and to identify the laws governing its use. They found that respondents' perceptions of the implications of legal restrictions were not always consistent with the actual legal language and that their perceptions may result in data use policies that are more stringent than what the law requires.[9]

The Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) has funded cross-agency partnerships with the Centers for Disease Control and Prevention, the Centers for Medicare & Medicaid Services (CMS), and the U.S. Food and Drug Administration to enhance the NDI-linked data available to researchers. One such study, "Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research," sought to add cause of death variables from the NDI to other large datasets, such as NCHS inpatient and emergency department claims data,[22] CMS Master Beneficiary Summary File and Chronic Conditions Warehouse files, Medicare and Medicaid decedent files, and Medicare and Medicaid research files.[23] The study's researchers outlined the steps necessary to apply for and obtain approval for linkage with the NDI and reported lessons learned;[24] they also created files for other researchers to use for PCOR projects and instructions on how to obtain them.[22]

In this current OS-PCORTF–funded project, we conducted a systematic search for statutes and regulations on the use of linkable mortality data for PCOR of each of the 57 mortality-reporting jurisdictions, as well as relevant logistical information (e.g., points of contact [POCs], application process, fee schedule) found on the jurisdictions' official websites. This information was then compiled into a database and an electronic compendium that PCOR investigators studying patient survival and mortality could use as a guide to determine which jurisdictions' data sources might be the most relevant to their research.

Purpose

The aim of this project was to develop a database of jurisdictions' policies governing linkable mortality data for research. This report summarizes the content of the database and describes the findings and lessons learned from compiling the database. We identify best practices to facilitate survival research via transparent policies and procedures around data acquisition and identify the jurisdictions that employ these practices. We also describe research use cases and illustrate how this database can be used by researchers and public health officials who are interested in conducting survival research.

Organization of Report

This report is organized into four chapters: (1) Introduction, (2) Methods (scope of the review, search strategy, and abstraction), (3) Overview of Findings (description of policies by category), and (4) Implications for Researchers. The report concludes with a summary and considerations for future research. Further details about our systematic search for statutes and regulations on the use of linkable mortality data for PCOR of each

of the 57 mortality-reporting jurisdictions and our codebook can be found in Appendix A and B, respectively. This report is also accompanied by an electronic compendium [LINK] that allows researchers to identify and filter jurisdictions with key characteristics and review source data for the policies.

CHAPTER 2. METHODS

Scope

The scope of this review covers formal policies (i.e., statutes and regulations), as well as administrative policies, processes, and application materials from the responsible agency or department in each of the 57 mortality-reporting jurisdictions.

Search Strategy

To identify relevant statutes and regulations, we searched all jurisdictions in WestLaw (a legal research database) under the "Statutes and Court Rules" and "Regulations" categories. Policies were returned in the search if they had at least one search term related to vital statistics or death records and at least one search term related to acceptable use of data (shown in Table 2.1). These terms were developed in consultation with a public health lawyer with experience using mortality data for research, the Assistant Secretary for Planning and Evaluation (ASPE), and a technical expert panel (TEP) consisting of experts from different federal agencies and with direct interest in the availability of linkable mortality data for PCOR. The full search string is presented in Appendix A.

Table 2.1. Search Strategy

Domain	Search Terms
Vital statistics or death records	"mortality data," "vital statistic," "death record!," "mortality record!," "death data," "vital record!," "mortality statistic!," "death statistic!"
Acceptable use	identif!, deidentif!, link!, permi!, authoriz!, reus!, restrict!, prohibit!, privacy, "institutional review board," IRB, NDI, "Death Index", disclos!, access!, research!

NOTE: An exclamation mark (!) indicates a wildcard character in the search term (e.g., identif! returns results for "identify" and "identification"). Terms enclosed in quotation marks require the exact words for a match.

We also searched each jurisdiction's website for relevant documentation regarding access to mortality data, including application forms and instructions. We began by visiting the jurisdiction's webpage for vital records. For jurisdictions where relevant documentation was not available at that location, we searched medical examiner and/or coroner's office and health department websites. Relevant webpages and online documents were searched by browsing site maps and evaluating hits from targeted searches using a website's search function (where available). Most abstracted websites were state government websites, but a small number were websites of third-party data vendors or state universities that collaborate with state governments on data management and access.

When little to no online information was available, or when only an email address was provided to request more information on obtaining mortality data, we reached out to the jurisdiction directly to request that information.

Screening Strategy

We screened all statutes, regulations, guidance documents, and related records returned by the search strategy described above. Two study team members redundantly screened a sample of documents according to a predefined set of inclusion and exclusion criteria and discussed any disagreements, clarifying inclusion criteria as needed. The final set of inclusion and exclusion criteria used are shown in Table 2.2.

Table 2.2. Inclusion and Exclusion Criteria

Type of Criteria	Criteria
Inclusion criteria	• Addresses release or use of individual-level data (not just aggregate) • Addresses death records or death data • Addresses mortality data with identifiers that would allow for record linkage or mortality data that could be linked to other datasets • Pertains to access to multiple records (as opposed to obtaining a copy of a single death record or death certificate)
Exclusion criteria	• Has been repealed or superseded by new policies • Related only to proposed legislation • Pertains to the use of death records in special investigations, criminal investigations, etc. • Pertains to the treatment of fetal death records • Pertains to the electronic death registration system • Pertains to death data that are not part of a vital statistics system (e.g., injury death reports)

Abstracting Policy Information

The study team developed a list of six policy dimensions through an initial scan of a sample of jurisdictions' vital statistics webpages and applications for access to mortality data, a review of the NDI data application, and consultation with ASPE, the TEP, and the public health lawyer. Policy dimensions and associated fields for abstraction are listed in Table 2.3. The team then developed a detailed codebook to guide abstraction of each policy document. Study team members redundantly coded a sample of policies with two reviewers coding each policy document. The team discussed their results, discrepancies in the redundant coding, and revised the coding guidance to clarify ambiguous cases, as needed. The team then singly coded the remaining policy documents and met regularly to bring questions to the group for discussion. Team members used DistillerSR® to screen and abstract policies. The final codebook is provided in Appendix B.

Table 2.3. Summary of Policy Dimensions

Policy Dimension	Summary of Data Fields
Data description	Available data elements, available years of data, preliminary data available, data completeness and gaps
Application and review process	POC, application process and timeline, fees, data formats
Legal, ethical, and privacy review requirements	Privacy or ethics review requirements, including whether Institutional Review Board (IRB) is specifically mentioned; documentation of proper use required; penalties for improper use
Permitted and disallowed use and users	Restrictions on the use of data and who is allowed to use data
Data protection requirements	Data safeguarding plan required, data elements considered personally identifiable information (PII), data storage and access requirements, data disposition plan requirements
Data linkage	Allowed linkage methods and existing linked datasets

Database Updates and Quality Assurance

After the initial search and abstraction that occurred March–July 2023, each jurisdiction's abstracted information was independently reviewed by a different member of the study team than the person who originally abstracted the information, and information was discussed by the team and corrected as needed. The database was also holistically reviewed and spot-checked by ASPE, a RAND quality assurance reviewer, the public health lawyer, and the TEP.

From December 2023 to June 2024, the study team then reviewed the website URLs in the initial database to determine whether the links were still active and whether there were any changes affecting the abstracted data. Where information had changed, the database was updated. Where links were no longer active, the team searched for revised URLs from the jurisdiction that contained the abstracted information and updated the database as necessary. The WestLaw search results were updated in March 2024 after searching for new policies not included in the initial search. These policies were then abstracted and incorporated. WestLaw was also used in June 2024 to identify statutes and regulations from the initial search with KeyCite's "red flag" status, meaning that at least part of the policy was repealed or superseded by a new policy. These policies were investigated and removed or updated as necessary.

In April 2024, the team sent the abstracted jurisdiction-level information to a vital statistics representative, identified by a member of the TEP, from each jurisdiction, offering them the opportunity to suggest additional information, sources, or corrections. The team incorporated this feedback into the database, as appropriate.

CHAPTER 3. OVERVIEW OF FINDINGS

Summary of Sources Identified

The final policy database includes 137 statutes, 80 regulations, and URLs for 190 websites.

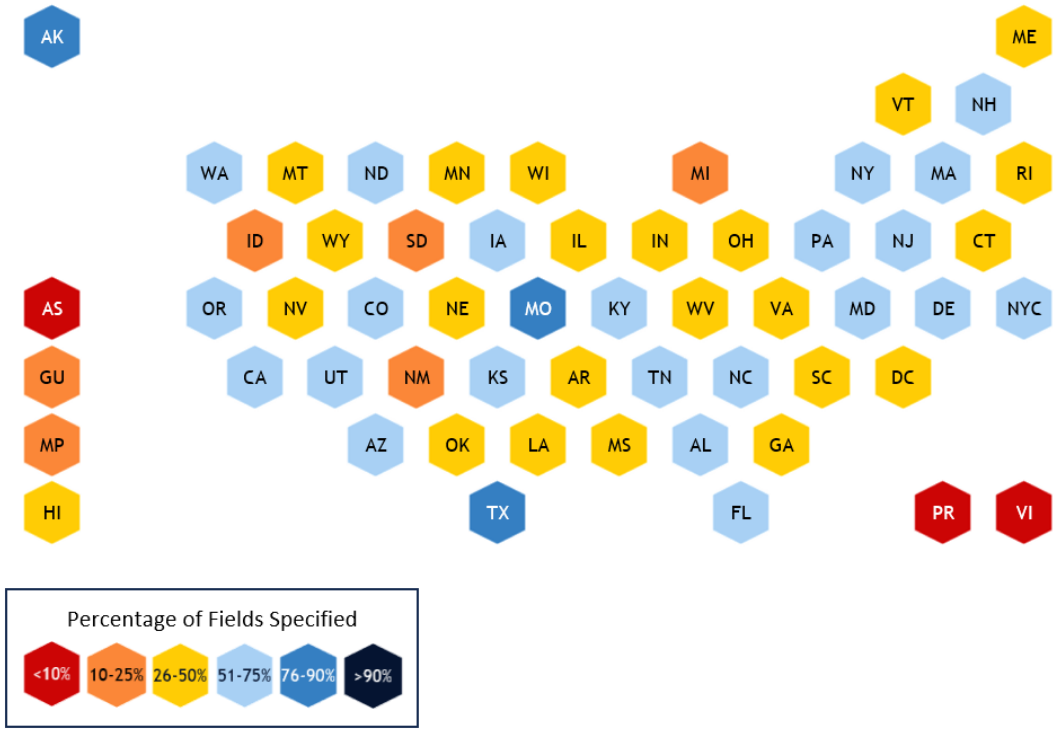
In our initial search of jurisdiction websites, described in the prior chapter, we found at least some application information from 24 jurisdictions. We contacted individuals from the 33 jurisdictions whose complete application materials could not be found online. Of these, we received responses from 15 jurisdictions (AS, CO, MA, MS, NC, ND, NY, OH, OK, SC, SD, UT, VA, VT, and WI). We were unable to ascertain information about the remaining 18 jurisdictions that did not respond.

In the second phase of the project, we contacted all 57 jurisdictions and asked them to review our findings and supply any corrections or additions. We sent one initial email with the jurisdiction's data attached in a table and two follow-up emails in the event of non-response to the initial request. We ultimately received feedback from 22 jurisdictions (AK, AL, AR, AZ, CO, GA, HI, IL, KS, KY, MA, MD, MO, NC, ND, NJ, NV, OK, PA, TX, UT, and WY), acknowledgement of receipt but no feedback from four (ID, MP, VT, and WA), and no response from the remaining 31 jurisdictions.

Summary of Abstracted Policy Data

Altogether we abstracted information on 35 data elements across six policy dimensions. There was substantial variation in the amount of information on policies we were able to find for each jurisdiction (Figure 3.1). We were able to abstract information for at least 75 percent of the 35 data fields for three jurisdictions (AK, MO, and TX). By contrast, we were able to extract information on less than 10 percent of the data fields for three jurisdictions (AS, PR, and VI).

Figure 3.1. Percentage of Data Elements with Specified Information, by Jurisdiction



Similarly, we found significant variability in how many jurisdictions provided information for each data element (Table 3.1). For example, we were able to identify a general POC name or entity for all but two jurisdictions. We were able to identify both POC information and a description of the permitted uses of mortality data in all but four jurisdictions. By contrast, very few jurisdictions specified any information on restrictions on which causes of death could be released or whether certain users were disallowed.

Table 3.1. Summary of Specified Information, by Data Element

Policy Dimension	Data Element	Specified Information, *N* (%)
Data description	Available data elements	32 (56%)
	Available years of data	17 (30%)
	Preliminary data available	14 (25%)
	Data completeness and gaps	14 (25%)
Application and review process	POC name or entity	55 (96%)
	POC contact information	52 (91%)
	Type of application process	51 (89%)
	Timeline from submitted application to data access	28 (49%)
	Fee description	42 (74%)
	Data format	29 (51%)
	Information about use case or purpose required	43 (75%)
Legal, ethical, and privacy review requirements	Privacy or ethics review requirements	34 (60%)
	IRB review and approval required	41 (72%)
	Documentation of proper use required	43 (75%)
	Penalties for improper data use or violations	23 (40%)
Permitted and disallowed use and users	Restrictions on the number of records	8 (14%)
	Restrictions on the cause of death	4 (7%)
	Other restrictions	8 (14%)
	Permitted uses	53 (93%)
	Disallowed uses	18 (32%)
	Allowed users	50 (88%)
	Disallowed users	4 (7%)
	Restrictions on the dissemination of results	34 (60%)
	Ongoing or open-ended access	17 (30%)
	Follow-back investigations	17 (30%)
Data protection requirements	Data safeguarding plan required	32 (56%)
	Data elements considered PII	23 (40%)
	Where must data be stored?	13 (23%)
	Where can data be accessed?	8 (14%)
	Is a disposition plan required?	33 (58%)
Data linkage	Can researchers link data themselves?	15 (26%)

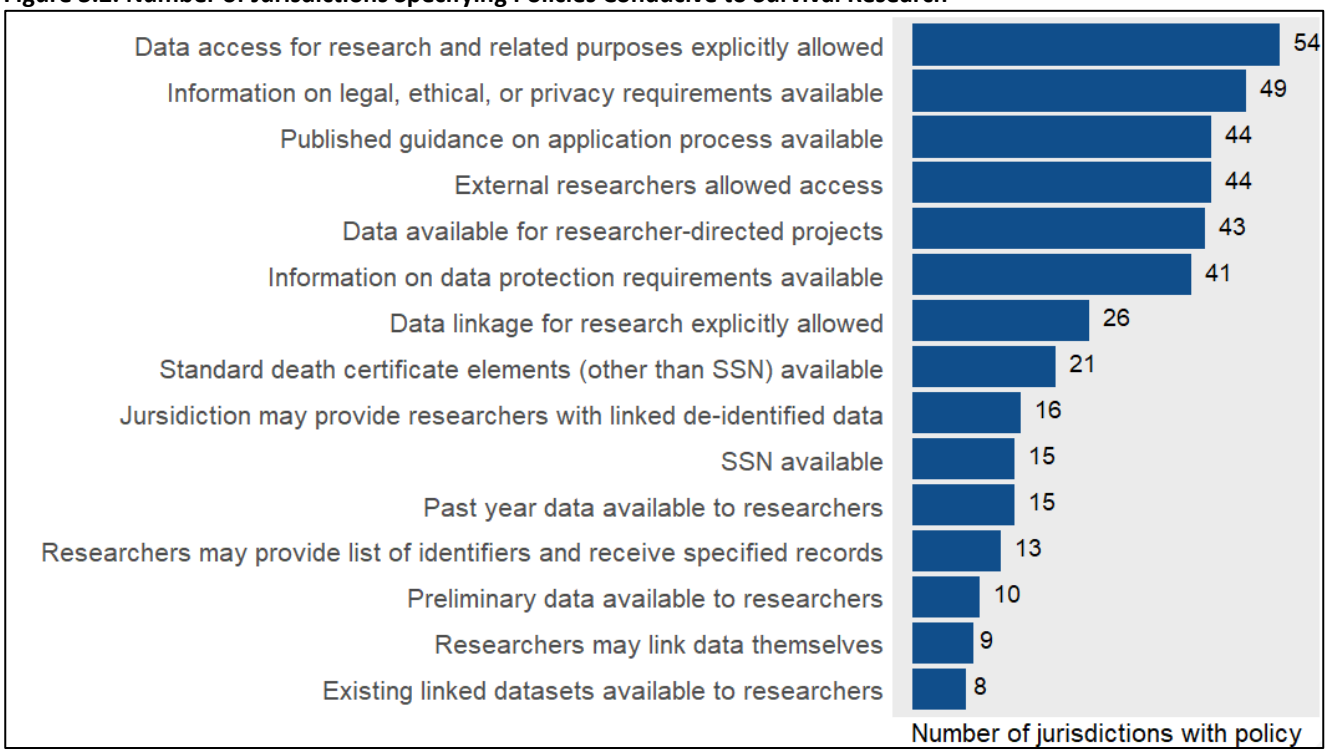
Policy Dimension	Data Element	Specified Information, *N* (%)
	Can the vital statistics office or another agency provide a list of names based on identifiers submitted by researchers?	12 (21%)
	Can the vital statistics office, another agency, or a third party do the linkage and provide researchers with the linked data?	18 (32%)
	Policies for additional data linkage	9 (16%)
	Summary of existing linked datasets	10 (18%)

We also reviewed the abstracted information and used it to identify whether jurisdictions had key features that may be important to researchers who want to use linked mortality data. These features are displayed in Figure 3.2. Nearly all jurisdictions (54) explicitly stated that mortality data could be used for research or related purposes either in statute, regulations, or other policy documents. However, we were only able to identify published guidance on how to access mortality data for research in 44 of these jurisdictions. Furthermore, we were only able to verify that these data were available to external researchers (e.g., not affiliated with the jurisdiction's government) in 44 jurisdictions and available for researcher-directed projects (i.e., the researcher can propose a project, as opposed to conducting research as a government contractor for a specific project defined by the jurisdiction) in 43 jurisdictions.

Just fewer than half (26) of jurisdictions explicitly noted that mortality data could be linked for research. Of these jurisdictions, most provided additional detail on how data could be linked. Most commonly, jurisdictions noted that they could link mortality data with other data sources and return the de-identified, linked data to researchers (16 jurisdictions). The next most commonly specified permitted way of linking data was for researchers to provide the jurisdiction with a list of identifiers and the jurisdiction could return the matching mortality records (13 jurisdictions). The least commonly specified permitted method was for jurisdictions to provide identifiers and allow researchers to create the linkages themselves (nine jurisdictions); five jurisdictions explicitly forbade researchers from creating linkages themselves. A few jurisdictions (eight) regularly linked mortality data with other data sources and made these data available to researchers.

Just fewer than half (21) of jurisdictions noted that standard death certificate elements other than SSNs were available to researchers. Of these, 15 noted that all standard death certificate elements could be requested by researchers. We note that this does not necessarily mean that SSNs would be provided to researchers for any study; only that the jurisdictions noted that standard death certificate elements were available and did not explicitly prohibit the release of SSNs to researchers. Most jurisdictions did not specify the most recent year of data available to researchers, but some (15) noted that the past year of data was available to researchers. In many jurisdictions, recent vital statistics data are classified as "preliminary" or "provisional" and subject to revision. Ten jurisdictions noted that these data were available to researchers.

Figure 3.2. Number of Jurisdictions Specifying Policies Conducive to Survival Research



NOTE: Bars represent the number of the 57 mortality-reporting jurisdictions that have policies explicitly allowing stated activities or making data available.

Details on Abstracted Data

Six common elements emerged in the data abstracted under each policy dimension. The commonalities are described in detail below. The full information abstracted for each jurisdiction with corresponding references is available in this report's accompanying electronic compendium [LINK].

Data Description

Within jurisdictions that described available variables in mortality data, most indicated that they provided standard death certificate data elements. In some cases, jurisdictions indicated specific data restrictions. Some jurisdictions did not include identifying variables, including SSN, as part of the available datasets (CT, NJ, OH, PA, RI, TX, VT, and WY). Other specified limitations were characteristics of the deaths, including time of death in one jurisdiction (TN) and cause of death in another (MT). In many jurisdictions, available variables were not specified or easily identified in public resources. In one state, resources noted that a data dictionary was available upon request (KS). In another jurisdiction, only variables available for matching were specified (NYC).

Among the 16 jurisdictions that specified years of data available, six jurisdictions indicated that death data were available starting prior to 1970 (GA, KS, MA, NC, PA, and WA). Excluding these, four jurisdictions indicated that death data were available starting between 1970 and 1980 (AK, CA, CO, and NYC), and the remaining jurisdictions indicated that specified available starting years of data were more recent (CT, KY, MN, ND, OR, and TX). The majority of jurisdictions did not specify the available years of data or only indicated the length of time that must elapse before death records become publicly available, which ranged from 25 to 75 years.

Some jurisdictions specified that preliminary data were available, sometimes referred to as "provisional." Such data included those that would be subject to amendment before the release of finalized data files. One jurisdiction specified that preliminary data were released as quarterly files, whereas all annual data were

considered final (WA). Most jurisdictions did not specify the availability of preliminary data.

When jurisdictions described data gaps or completeness, they included descriptions of when death certificate forms were changed (CO, GA, MO, NJ, and TX) or provided a general statement noting that available data elements vary across years and/or datasets (AZ, CA, CT, NC, ND, NYC, OR, and WA).

Application and Review Process

The study team identified some level of detail on the application and review process for acquiring mortality data in all but six jurisdictions (AS, ID, MI, NM, PR, and VI; a representative from AS told us that there was no process in place for sharing mortality data with researchers). However, as noted above, some of this information is only available by directly requesting application materials from the jurisdiction. All jurisdictions that specified an application process noted an initial data request, many via standard application, followed by jurisdiction review. In some cases, approval of a screening application is followed by a request for more-detailed materials from the applicant (e.g., TX). The precise process for application and review is unique to each jurisdiction with varied intensity, but many processes share common elements. Applications commonly require researcher information and credentials, a description of the study and associated need for the data, specification of requested variables and years of data, information on prior or current IRB activity, and DUAs. Jurisdiction resources commonly note that clarification or additional information may be requested following the initial review of application materials. The parties responsible for review and approval of data requests vary from jurisdiction to jurisdiction.

Most jurisdictions provided instructions on the elements to include in applications. The required elements of applications commonly included: (1) study purpose and objectives; (2) hypotheses or research questions; (3) use of requested data (including methodologies, analyses, and anticipated results); (4) anticipated benefits to society or specific populations; (5) justification for each data variable requested; (6) study protocol, including study design, time frame, and target population; and (7) plans for publication or release of study findings. In three jurisdictions, an explanation of how study findings would specifically help the residents of the jurisdiction was requested (IA, NH, and NY).

Jurisdictions that provided a timeline for the application process gave a wide range of expected processing times, commonly noting that time frames depend on the nature and complexity of the request and availability of staff to fulfill the request. Published timelines ranged from weeks to several months and were sometimes broken down by stage of the application process (e.g., KS, MD, and WA).

Jurisdictions used a variety of fee structures for the use of mortality data. Types of fees included per dataset fees, per record fees, per hour fees for data preparation and or analysis, annual ongoing charges, and flat fees. Fees typically varied by type of service or staff type, with higher fees associated with more-skilled labor. For hourly fees, rates ranged from a low of $15 per hour for administrative staff (SC) up to $125 per hour for complex requests (WY). Sixteen jurisdictions' websites or legislation indicated that fees applied, but they did not specify the amount of these fees (AL, AR, DC, DE, ID, IL, IN, MA, ME, MS, NH, NM, NYC, OK, TX, VA, and WI). Three jurisdictions did not charge fees (MA, ND, and NJ).

Most jurisdictions that specified available file formats accepted requests for digital transfers (e.g., using Secure File Transfer Protocol [SFTP]). Eleven jurisdictions specified that digital and hard copy (e.g., compact disc, flash drive) formats were available (AR, CA, CT, FL, KS, MO, MS, NH, NY, TN, and WY). One jurisdiction specified availability of only hard copy formats (NC).

Legal, Ethical, and Privacy Review Requirements

Most jurisdictions specified that IRB review is required to access mortality data, whether from the applicant's own institution, a jurisdictional IRB, or both. Among most jurisdictions where the IRB review requirement was

not specified (AS, CT, GA, GU, ID, IN, MA, MN, MI, MP, MS, NM, NV, OH, PR, SD, VI, and WV), several required review from other bodies, such as a data council (MP), a legal reviewer (MN and MS), or the state registrar (NM). In some jurisdictions, additional review may be required by a privacy board or its equivalent (AR, CA, DE, KY, MD, TN, and WI).

A data use agreement (DUA) or other attestation or assurance of appropriate data use is also required by most jurisdictions to obtain mortality data. Only one jurisdiction specified that a DUA was not required (NC), while some jurisdictions specified that DUA requirements depended on the data requested (KY, MI, and UT). Among jurisdictions where DUAs were not specifically mentioned but where other attestations or assurances were, common alternatives included signed confidentiality agreements (AZ, HI, MO, NM, and NY) and attestations of proper data use on applications themselves (DE, NE, and TX).

Among jurisdictions that specify penalties for the improper use of data, penalties fell into several categories. Some states specified fines (AZ, CA, DC, GA, MO, PA, WV, and WY), ranging from no more than $100 (WY) to up to $12,500 (DC). Jurisdictions also described other financial consequences, such as civil liability. Criminal liability and imprisonment were specifically mentioned in many jurisdictions (CA, DC, GA, DC, MO, NE, NH, NM, OR, PA, WA, and WY). Beyond these specific penalties, violations could lead to the termination of data access and future access to data, termination of research agreements, and notifications to oversight bodies, such as IRBs and health departments.

Permitted and Disallowed Use and Users

No jurisdictions specified a maximum number of records that they would release, but a few (DE, NJ, NYC, TX, and WY) noted that they would release only the minimum number of records necessary to answer the proposed research question, without specifying how that minimum number would be determined. No jurisdictions put restrictions on which causes of death they would release (e.g., drug overdose, suicide), provided that they released the cause of death variable to researchers. Other restrictions mentioned by jurisdictions included restrictions on how long researchers could access the data (MO and NY), not releasing individual-level data if aggregated data could meet research needs (MD), or sharing only record-level information with the consent of the immediate family of the deceased (MS).

As noted above, nearly all jurisdictions explicitly allowed mortality data to be used for research or closely related purposes, such as “statistical,” “epidemiological,” “public health,” “surveillance,” “medical,” “scientific,” “health care operations,” or “quality improvement” purposes. Furthermore, nearly all jurisdictions specified research as an allowed use of mortality data in statute. Of jurisdictions that noted disallowed uses, most specified commercial use. Other specified disallowed uses included causing embarrassment to family members (HI), reconciliation of records (KY), or speculative purposes (GA, NYC, OR, and SC).

Most jurisdictions specified allowed users. Of these jurisdictions, allowed users specified by more than half of jurisdictions included specified government agencies and researchers. Other specified allowed users included “public or private agencies” (DC, IA, IL, KY, MD, MI, MN, MO, NM, NV, RI, SC, SD, and VA), “government contractors” (IN and OR), “individuals” (CA, KS, MP, NM, WY, and WA), and “health care providers” (HI, KY, MD, ND, OK, OR, UT, and WA). While these jurisdictions specified users of mortality data, it was often not clear whether all users could access individual records and all available variables. Only four jurisdictions specified disallowed users, and all of these jurisdictions specified “commercial firms or agencies” (AL, NH, UT, and WV).

Several jurisdictions specified restrictions on the dissemination of research results. Frequently specified restrictions included required approval of dissemination products by the department or agency (IA, MO, RI, and UT), notification of publishing (MD, NC, NJ, and TX), or acknowledgement of the department or agency as the source of research data (CO, MD, PA, TN, and TX). Several jurisdictions specified that releasing PII or small cell counts (AK, AZ, CO, DC, FL, HI, IL, KY, MO, ND, NJ, RI, and SC) is prohibited. One jurisdiction (MO) also

required approval from the jurisdiction to release information on individual health care providers.

Most jurisdictions did not specify where open-ended access to data was allowed. Of the jurisdictions that did specify policies on open-ended or ongoing access, no jurisdictions allowed open-ended access to mortality data, most jurisdictions (AK, AZ, CA, CO, DE, KY, ME, MO, ND, NY, OK, OR, RI, TX, UT, and WA) allowed extensions to data access with a renewal application, and a few jurisdictions (IA, KS, and NC) required a completely new application to extend access to the data. Most jurisdictions also did not specify whether follow-back investigations were permitted. Of the jurisdictions that did specify policies on follow-back investigations, most (AK, AL, AZ, HI, IA, KS, LA, MD, ME, MO, NC, OR, PA, TX, and WA) allowed them with a special application, though many of these jurisdictions emphasized that doing so must be determined essential to conducting the research. Two jurisdictions (CO and ND) did not permit follow-back investigations for external researchers, though CO noted it may be possible for State Health Department–authorized public health investigations. One jurisdiction (FL) noted that "intrusive" follow-back was not permitted, but it was unclear if follow-back was prohibited under all circumstances.

Data Protection Requirements

Most jurisdictions that required a DUA or other attestation to obtain data also required that applicants submit a data safeguarding plan (DSP) to explain how they would ensure the security of the data they receive. Most commonly, requirements for DSPs require researchers to describe administrative, technical, and physical safeguards. Nearly half of jurisdictions noted specifically which variables were considered PII. These variables commonly included names, birth dates, death dates, SSN or other identification numbers, and small geographic units (the definition of small geographic unit varied by jurisdiction).

While most jurisdictions had requirements in the DUA or other attestation that data be kept secure, most jurisdictions did not specify where data should be stored. Some jurisdictions specified data storage requirements (AK, AL, CA, FL, IA, KS, MD, NJ, NY, NYC, OR, TX, and UT) or did not have specified requirements but noted that the storage location must be approved by the jurisdiction (CO and ND). Most jurisdictions had requirements in the DUA that only authorized individuals could access the data. Beyond this requirement, most jurisdictions did not specify how data could be accessed. Some jurisdictions noted that an access plan must be approved by the jurisdiction (AL, CO, KS, and UT). The few jurisdictions that specified access requirements (AK, AZ, CA, FL, IA, and MD) included access through password-protected workstations, automated screen time-outs, and security log reviews. One jurisdiction (DE) required any analysis using PII to be analyzed on-site at the health department.

Most jurisdictions required a data disposition plan. Some jurisdictions included an attestation that data would be destroyed at the conclusion of the project as part of the DUA (AK, KS, and UT), others described how data destruction must occur (CA, DE, ME, and TN), others required descriptions of how and when the data would be destroyed as part of the application (AL, CO, FL, KS, MT, NY, NYC, OR, SC, SD, VA, and WA), and others required that a certification of data destruction be returned at the end of the project (AZ, IA, IL, MO, NH, RI, and TX).

Data Linkage

Within jurisdictions explicitly allowing data linkage, there was considerable variation in the ways in which data could be linked. Across all methods of linkage, plans to link typically needed to be described in an application for the data and approved by the jurisdiction. For all jurisdictions allowing researchers to link data themselves, researchers needed to complete some type of application and/or DUA, typically outlining the linkage variables, the steps for data linkage, and the uses of the linked data. One jurisdiction (OK) only allowed researchers to conduct linkage on-site at the state health department's facilities. Most jurisdictions did not specify the variables or methods that could be used for linkage. However, one jurisdiction (TX) noted that SSN was not available for researchers conducting linkages themselves but could be used if the health department

conducted the linkages and returned de-identified data to researchers. Pennsylvania noted that, when researchers provided identifiers to receive matching death records, last name, first name, date of birth, and SSN (or at least the last four digits of SSN) are needed for accurate linkage results. Pennsylvania also noted that address and premarital name could be used to "resolve questionable matches."

No jurisdictions explicitly disallowed all types of data linkage. A few jurisdictions (CO, MS, PA, and VA) explicitly disallowed researchers to link data themselves but could accommodate requests for the department to link the data and return de-identified data to the researchers. Some jurisdictions did not explicitly allow researchers to conduct data linkage but indicated that researchers could request identifying information that could potentially be used for data linkage (CT, HI, and IA). Some jurisdictions' data applications had questions about data linkage, but it was unclear whether researchers could perform the data linkages themselves (DC, NC, NJ, OH, RI, SD, and WA).

A few states routinely link mortality data with other datasets and, under some circumstances, make these linked data available to researchers. For example, California has death data linked with emergency department and hospital discharge data, Kansas has linked birth and death data, and Missouri has a birth defect registry linked with birth and death records. Massachusetts has a large data warehouse that includes 35 linked datasets, including birth and death records, health insurance claims, prescription drug–monitoring data, and criminal justice data. However, these linked data are typically only available to Massachusetts government employees or contractors or to researchers responding to specific notices of opportunity.

CHAPTER 4. IMPLICATIONS FOR RESEARCHERS

Mortality data can be a useful source for public health research and PCOR, and most jurisdictions note that these data are available to researchers. However, the utility of mortality data depends on several features, including how the data may be used, which data are available, application and access procedures for the data, and how the data may be linked to other sources. In this chapter, we describe some features of mortality data policies that are important for researchers and highlight jurisdictions with model policies that include these key features.

Key Features of Mortality Data Policies for Researchers

This section describes important features of mortality data policies for researchers (summarized in Table 4.1), determined based on a review of the results in Chapter 3 and on input from ASPE and mortality researchers on the study team. Features are characterized as "essential" if they are necessary for most PCOR-related projects and "preferred" if they are necessary for certain types of PCOR projects or could reduce barriers for researchers. The electronic compendium accompanying this report allows researchers to filter on all the essential features and select preferred features to identify jurisdictions that meet their research needs [LINK].

Table 4.1. Key Features of Mortality Data Policies for Researchers

Policy Dimension	Essential Features	Preferred Features
Data description	• Standard death certificate elements (other than SSN) available to researchers • Past year data available to researchers	• SSN available to researchers • Preliminary data available to researchers
Application and review process	• Published guidance on application process available	• Timely data application approval and data delivery • Low fees
Legal, ethical, and privacy review requirements	• Information on legal, ethical, or privacy review requirements available	• Flexibility in allowing a researcher's or a jurisdiction's IRB to approve the study
Permitted and disallowed use and users	• Data access for research and related purposes explicitly allowed • Data available for researcher-directed projects	• Data explicitly allowed to be used for public health purposes, quality improvement, and surveillance • Data available to external researchers
Data protection requirements	• Information on data protection requirements available	• On-site analysis of data not required • Policies that provide both sufficient data protection and a level of access necessary for research
Data linkage	• Data linkage for research explicitly allowed	• Researchers can link data themselves. • Researchers may provide list of identifiers and receive specified records. • Jurisdiction may provide researchers with linked de-identified data.

Data Description

The availability of standard death certificate elements, particularly cause of death and identifiers for linkage, are essential for many types of PCOR. Cause of death is important to measuring mortality outcomes, while identifiers are needed to match exposures or participants from primary data collection to death records. Some jurisdictions have restrictions on releasing SSNs, even for deceased individuals, due to fraud concerns. However, if proper protections are in place, SSNs could be used to improve the specificity of data linkage, because linking on only such identifiers as name and birth date can lead to false matches. Using SSN in combination with other identifiers to link data can also increase the sensitivity of matches in cases where some identifiers are incorrect (e.g., misspelled names) or missing.

Application and Review Process

We identified many different types of application and review processes that allow researchers to obtain mortality data. While there is no one preferred application and review process, researchers need documentation of these processes and materials to be available. For example, publishing data applications, DUA templates, data dictionaries, fees, and timelines would help researchers greatly as they develop their research plans. Ideally, such documents would be publicly available (i.e., posted on a website). Timely access to data (e.g., less than three months from data application to delivery) is preferred, as research is often time-sensitive (e.g., producing actionable evidence on the COVID-19 pandemic) and must often be completed within a specified time frame due to funding restrictions. Keeping data access fees low (e.g., no more than the cost to the jurisdiction of processing the request) is also preferred, as affordability encourages a greater volume and diversity of research.

Legal, Ethical, and Privacy Review Requirements

Jurisdictions differed in their published policies on legal, ethical, and privacy review requirements. While there is no one preferred approach, researchers need transparent requirements that balance the privacy of deceased individuals and their families with practical considerations for researchers. A preferred feature is to allow either the researcher's or the jurisdiction's IRB to approve the study in lieu of requiring additional approval from the jurisdiction.

Permitted and Disallowed Use and Users

An explicit statement allowing mortality data to be used for research is characterized as essential. In some cases, it may be useful for health researchers and practitioners to use data for other related purposes, such as public health, surveillance, or quality improvement, and statements explicitly allowing these activities are characterized as preferred. Making data available to researchers for researcher-directed work (i.e., not as part of a specific contract or narrow call for research by the jurisdiction) expands opportunities for novel public health research and discoveries. While having a jurisdiction's government as a research partner can be very valuable, when feasible, it is preferred that external researchers (e.g., not employed by the jurisdiction's government) can access the data to increase the number and diversity of research projects using mortality data.

Data Protection Requirements

Jurisdictions described a variety of data protection requirements. In general, it is ideal when jurisdictions' data protection requirements are transparent, meaning that requirements are readily available to researchers. It is also preferred that policies provide adequate data protection and are feasible for researchers to implement. On-site data analysis requirements can be a significant barrier for many researchers.

Data Linkage

Data linkage is critical for many PCOR studies, and allowing certain types of linkage may also be preferred for specific research use cases. These types of linkages and research use cases are as follows:

- Researchers may link data themselves: In this use case, the jurisdiction provides researchers with mortality data, including identifiers that allow researchers to link data themselves. This ability may be desirable for several reasons. For example, researchers could have more flexibility over the linkage timeline and methods used. They may be able to perform linkages more quickly than the jurisdiction or use complex probabilistic linkage methods that the jurisdiction does not have the capacity to perform. Researchers may also be able to link the data to other datasets to which the jurisdiction does not have access (e.g., primary data collected by the researchers). This method also puts less strain on jurisdictions' time and resources. However, some jurisdictions do not permit the release of specific identifiers (e.g., SSN) or identifiers more broadly, either by statute or informal policy, even for deceased individuals. Jurisdictions may find this method less preferred because they have less control over how the mortality data are used in these circumstances.
- Researchers may provide a list of identifiers and receive specified records: In this use case, researchers provide the jurisdiction with a list of identifiers and then the jurisdiction returns death records for all individuals matching those identifiers. This method is useful in cases where researchers have conducted an earlier clinical trial or other primary data collection effort, and they need to assess longer-term mortality data of the participants. However, researchers are limited in the linkage methods they may use. For example, in some cases, the jurisdiction will only do deterministic linkage, and researchers may not receive a unique identifier, such as an SSN. Jurisdictions may prefer this method because it allows them to release the minimum number of records necessary to conduct the research.
- Jurisdiction may provide researchers with linked de-identified data: In this use case, the jurisdiction links mortality data with other datasets, removes identifiers used in the linkage, and then returns the linked de-identified data to researchers. This method is useful in cases where the datasets to be linked are owned by the jurisdiction. For example, a state health department may be the steward of the state's cancer registry and mortality data. Jurisdictions may prefer this method, as they can support research without releasing PII. However, this method does place additional burden on jurisdictions and limits researchers to datasets owned by the jurisdiction.

Model Jurisdictions Meeting Essential Policy Criteria

We identified five jurisdictions (KY, NC, OR, PA, and WA) that met all the essential policy criteria and some of the preferred criteria:

- All five jurisdictions made standard death certificate data elements other than SSN available to researchers and had data from the past year available. Two of these jurisdictions (KY and WA) may also allow SSN to be used for some research projects, and two jurisdictions (NC and WA) define preliminary data and make them available to researchers.
- All five jurisdictions had at least some published guidance on their application and review processes, though the level of detail available varied, and some information was only available through email correspondence with the jurisdiction. Three of the five jurisdictions (OR, PA, and WA) had data applications available online, one jurisdiction had information about the timeline of data requests and a data dictionary available online (WA), and four jurisdictions provided specific information about data fees (OR, NC, PA, and WA). Where specified, fees were generally modest and dependent on file size and processing time. However, one jurisdiction's (WA's) higher fees could be cost-prohibitive for some research projects, including a $3,000 fee for a new IRB review, $1,500 for a renewal, $175–$350 for each data file, and $100–$200 per hour of analysis time.
- All five jurisdictions provided at least some guidance on legal, ethical, and privacy review

requirements, though this information also varied considerably in terms of publicly available details. All five jurisdictions require IRB approval for at least some types of studies using mortality data. For studies in which IRB approval is required, two jurisdictions (KY and WA) require approval by the jurisdiction's IRB, one jurisdiction (OR) requires approval by the researcher's IRB, one jurisdiction (PA) provides an option to use either the jurisdiction's or researcher's IRB, and one jurisdiction (NC) does not specify which IRB is required to approve the study.

- All five jurisdictions explicitly allowed data to be used for research, including for researcher-directed projects. All five jurisdictions also allowed data to be used by external researchers. One jurisdiction (WA) also specified that the data could be used for "public health purposes."
- All five jurisdictions provided some information on data protection requirements. For example, all jurisdictions require submission of a data safeguarding plan, four jurisdictions (KY, OR, PA, and WA) require data disposition plans, and one jurisdiction (OR) specifies data storage requirements. None of the five jurisdictions required on-site analysis of data.
- All five jurisdictions allowed individual-level record linkage for research. Two jurisdictions (KY and OR) explicitly allowed researchers to link data themselves, two jurisdictions (PA and WA) noted that researchers may provide a list of identifiers and receive the specified linked records back, and three jurisdictions (NC, PA, and WA) noted that the jurisdiction could link the data and provide researchers with linked de-identified data.

Additional detail on these five jurisdictions is provided in the accompanying electronic compendium [LINK].

CHAPTER 5. REPORT SUMMARY AND CONSIDERATIONS

In this report, we examined statutes, guidance, and other publicly available policies governing the use of linkable mortality data for PCOR in each of the United States' 57 mortality-reporting jurisdictions. This review was supplemented by outreach to the jurisdictions. We abstracted information on the following policy dimensions:

- data description
- application and review process
- legal, ethical, and privacy review requirements
- permitted and disallowed use and users
- data protection requirements
- data linkage.

Almost all jurisdictions explicitly allowed mortality data to be used for research, but the amount of information we could find on these policy dimensions varied greatly by jurisdiction. Some jurisdictions provided detailed information on all six of these dimensions and publicly posted such materials as applications for data, timelines, fee schedules for acquiring data, and data use guides. Other jurisdictions provided essentially no information on their policies governing the use of mortality data for research.

After reviewing the available policy information, we identified five model jurisdictions (KY, NC, OR, PA, and WA) with policies facilitating the use of mortality data for PCOR. While these jurisdictions have several policy features that may be desirable for PCOR, researchers may still find that other jurisdictions can provide data meeting the needs of their specific projects. These researchers can explore the accompanying electronic data compendium [LINK] to explore details on policies in jurisdictions of interest or search for jurisdictions that have desired policy characteristics. Additionally, while we were unable to identify information on some policy dimensions, jurisdictions may have additional internal policies or practices. Researchers may still be able to successfully use data from these jurisdictions, particularly if they have a research partner from the jurisdiction.

Although there are barriers to using the NDI, researchers who are interested in studying survival at the national level or among individuals across several states would likely find that option to be less burdensome than requesting data from multiple jurisdictions. However, those researchers who require only a subset of data from one or two jurisdictions may find that requesting data directly from individual jurisdictions is a more cost-effective and less burdensome approach that may allow for more flexibility in linkage.

Policymakers might consider how to make jurisdiction-level mortality data more accessible to researchers in light of the potential scientific and public health gains that could be made. In general, we found that most jurisdictions explicitly allowed research in statute, and there were few legal restrictions on the use of mortality data for research. Where legal restrictions existed, they generally involved limitations on the use of SSN or the public release of PII. Given these findings, jurisdictions may consider emulating the policies of the five model jurisdictions in cases where public health departments or vital record offices may set their own policies.

Death records from jurisdictions are a vital source of data for PCOR. This study shows that it is feasible to conduct research using mortality data linked with other data sources in many jurisdictions. However, improved transparency on data access policies would facilitate important research.

APPENDIX A. SEARCH DETAILS

The complete search string is as follows:

(mortality data" OR "vital statistic" OR "death record!" OR "mortality record!" OR "death data" OR "vital record!" OR "mortality statistic!" OR "death statistic!") /50 (identif! OR deidentif! OR link! OR permi! OR authoriz! OR reus! OR restrict! OR prohibit! OR "privacy" OR "institutional review board" OR "IRB" OR "NDI" OR disclos! OR access! OR research! OR "Death Index")

Note that an exclamation mark (!) indicates a wildcard character (e.g., identif! returns results for "identify" and "identification") and /50 indicates that a word from the first group must be within 50 words of the first set.

APPENDIX B. CODEBOOK

Field Name	Data Format/Response Options	Description and Notes
Available data elements	• Standard death certificate data elements • Limited set of data elements (information describing what is missing or redacted—e.g., no SSN, identifiers) • Not specified	Data elements available to researchers. Standard death certificate elements are defined as those available on the 2003 U.S. Standard Certificate of Death.[1] For the second option, "limited set of data elements" is followed by a description of which variables from the standard death certificate data elements are not included in parentheses.
Description of available years of data (range of available data, etc.)	Text	Field includes information about all available years of data: most recent available, most recent complete available, first available, and year after death during which data becomes publicly available. For websites, this might be a date range. For policies, it will generally be a qualitative description (e.g., data might be made publicly available for deaths occurring more than 100 years ago; data must be updated quarterly). If no information or date range is given, "Not specified."
Preliminary data available (describe)	Text	If available data are considered preliminary (i.e., deaths for recent years are not considered complete), describe which years are considered preliminary. If preliminary data are explicitly not available to researchers, "Preliminary data not available to researchers." If no information on preliminary data is given, "Not specified."
Data completeness and gaps (describe)	Text	Description of any data gaps or details not described previously (e.g., variables missing for specific years, counties with missing data)
POC name or entity (free text)	Text	Name, role, or department responsible for death record data. If a specific name, role, or department for research requests is available, this contact will be listed. Otherwise, the department, name, or role responsible for death record data more broadly will be listed.
POC contact information (free text)	Text	Contact information for death record research requests, if available; otherwise, contact information for death records more broadly. Email, phone number, and address, where available, are included.
Type of process (free text)	Text	Summary of process, including who may review application, application steps, etc.
Timeline from submitted application to data access (describe)	Text	Description of any timeline information, number of months for review, etc.
Fee description (who must pay, fee amount or structure)	Text	Description of fee amounts and who must pay fees

Field Name	Data Format/Response Options	Description and Notes
Data format (digital [SFTP, etc.], hard copy [compact disc, hard drive, etc.])	• Not specified • Digital (secure file transfer, etc.) • Hard copy (compact disc, hard drive, etc.) • Other (describe)	Multiple options may be checked. Descriptions of specific data formats are included in parentheses after each option.
Description of information about use case or purpose required (free text)	Text	Description of information on use case or purpose that the data applicant must provide
Description of legal, ethical, and privacy review requirements (free text)	Text	General legal, ethical, and privacy review requirements (not IRB–specific). Includes information on who does the review and what the review entails.
IRB review and approval required	• Not specified • Yes (describe)	Only select yes if the term "Institutional Review Board" or "IRB" is specifically mentioned. Other types of review should be listed in the prior field. If "Yes," describe IRB process (which IRB needs to approve [e.g., researchers' institution, jurisdiction's institution, or both], who needs approval).
Documentation of proper data use required	• Not specified • No • Yes, DUA specifically mentioned • Yes, other attestation or assurance mentioned (describe)	If text specifies "data sharing agreement" or "DUA," then "Yes, DUA specifically mentioned" is recorded. If another type of attestation or assurance is mentioned, then "Yes, other attestation or assurance mentioned" option is selected and described in the parentheses that follow this option.
Penalties for improper data use, violations, and falsification (describe)	Text	A description of the penalties (e.g., fines, jail time) for data use violations or for falsifying information on an application for data. Penalties for falsifying information on a death certificate are not considered here.
Restrictions on the number of records (number or unrestricted) (describe or not specified)	• Not specified • No • Yes (describe)	Describes restrictions on the number of records that may be obtained.
Restrictions on the cause of death (describe or not specified)	• Not specified • No • Yes (describe)	Description of specific causes of death (e.g., overdose, suicide, or other specific causes of death) that are restricted or redacted in data access requests
Describe other restrictions (text)	Text	Any other restrictions on the availability of data for researchers that are not previously defined
Permitted uses	Text	Permitted uses of data, including exact language used
Disallowed uses (research, surveillance, quality improvement, reconciliation, commercial, other, and not specified)	Text	Disallowed uses of data, including exact language used

Field Name	Data Format/Response Options	Description and Notes
Allowed users	• Not specified • List (e.g., state government, federal government, researchers, nonprofit, commercial)	List of allowed data users specifically mentioned
Disallowed users	• Not specified • List (state government, federal government, researchers, nonprofit, commercial)	List of disallowed data users specifically mentioned
Restrictions on dissemination of research results	• Not specified • No • Yes (describe)	Restrictions on the dissemination of research results using identifiable or linked data. Examples: Description of dissemination activities required in application, review by jurisdiction required, no reporting of cell sizes of <10, no reporting of identifiers, etc.
Ongoing or open-ended access	• Not specified • Not permitted (describe) • Permitted (describe) • Permitted with special application (describe and specify limited or unlimited, e.g., limited number of times that user can reapply)	Not permitted = no extensions granted, completely new data application needed. Permitted = no end date by which researchers must stop using data. Permitted with special application = an application is needed to extend data use for original project or stated purpose; include information describing limits, lack of limit, etc.
Follow-back investigations	• Not specified • Not permitted • Permitted (describe) • Permitted with special application (describe)	Follow-back investigations = contacting individuals identified through data analysis or following up on specific individuals to obtain more information (e.g., following up with next of kin, getting more information on deceased's hospital records). Not permitted = not permitted under any circumstances. Permitted = allowed once access to data is granted. Permitted with special application = need a separate application for a follow-back investigation.
Data safeguarding plan required?	• Not specified • No • Yes (describe)	Describes measures put in place to protect data from unintentional release or disclosure.
Data elements considered personally identifiable information (PII) (list)	Text	List of data elements specified by the source as being PII. If no elements are specifically defined as PII, "Not specified."
Where must data be stored?	Text	Describes where data are required to be stored (e.g., encrypted hard drive, locked office)
Where can data be accessed?	Text	Describes where data are required to be analyzed (e.g., data center, cold room)
Is a disposition plan required?	• Not specified • No • Yes (describe)	Describes whether researchers must include in their application a plan to destroy the data after the project is completed

Field Name	Data Format/Response Options	Description and Notes
Can researchers link data themselves (i.e., vital statistics office or another agency provides researchers with identifiers that researchers can use to do linkages)?	• Not specified • No • Yes (describe identifiers available to the researcher to conduct the linkages)	Are researchers both able to access necessary identifiers and allowed to link data themselves? Identifiers may include SSN, or a combination of name and other variables, such as birth date, death date, etc.
Can researchers provide a list of identifiers to the vital statistics office or another agency and receive back a list of people with those identifiers?	• Not specified • No • Yes (describe identifiers that must be provided by the researcher)	Examples: Researchers must provide SSN; researchers must provide either SSN *or* birth date, death date, first name, and last name.
Can the vital statistics office, another agency, or a third party do the linkage and provide the researcher with the linked data (i.e., researchers do not have identifiers for linkages)?	• Not specified • No • Yes (describe identifiers used to conduct linkages)	Example: The vital statistics office links death data with another data source using SSNs, then gives the linked data file with SSNs removed to researchers.
Summary of existing linked datasets	Text	Summary of existing datasets that are already linked to death records that researchers can request (e.g., cancer registry linked with death certificate data; all-payer claims data linked with death certificate data). Only include data linked at the individual level (e.g., publicly available unemployment data at the county level linked by county would not be included).

ABBREVIATIONS

ASPE	Assistant Secretary for Planning and Evaluation
COVID-19	coronavirus disease 2019
DUA	data use agreement
IRB	Institutional Review Board
NCHS	National Center for Health Statistics
NDI	National Death Index
PCOR	patient-centered outcomes research
PII	personally identifiable information
POC	point of contact
SFTP	Secure File Transfer Protocol
SSN	Social Security number
TEP	technical expert panel

REFERENCES

1. National Center for Health Statistics, "U.S. Standard Certificate of Death," November 2003. As of July 17, 2024: https://www.cdc.gov/nchs/data/dvs/death11-03final-acc.pdf
2. Parpia, Alyssa S., Isabel Martinez, Abdulrahman M. El-Sayed, Chad R. Wells, Lindsey Myers, Jeffrey Duncan, Jim Collins, Meagan C. Fitzpatrick, Alison P. Galvani, and Abhishek Pandey, "Racial Disparities in COVID-19 Mortality Across Michigan, United States," *EClinicalMedicine*, Vol. 33, February 26, 2021.
3. Patient-Centered Outcomes Research Institute, "Establishing the Definition of Patient-Centered Outcomes Research," webpage, updated July 15, 2014. As of July 25, 2024: https://www.pcori.org/research/about-our-research/research-we-support/establishing-definition-patient-centered-outcomes-research
4. Lund, Jennifer L., Jenna Cacciatore, Rachel Tylock, I-Hsuan Su, Saloni Sharma, Sharon Peacock Hinton, Sabirah Smith, Molly A. Nowels, Xiaomeng Chen, Paul R. Duberstein, et al., "Development and Evaluation of a Multisource Approach to Extend Mortality Follow-Up for Older Adults with Advanced Cancer Enrolled in Randomized Trials," *JCO Clinical Cancer Informatics*, Vol. 8, 2024.
5. Mumma, Michael T., Sarah S. Cohen, Jennifer L. Sirko, Elizabeth D. Ellis, and John D. Boice, Jr., "Obtaining Vital Status and Cause of Death on a Million Persons," *International Journal of Radiation Biology*, Vol. 98, No. 4, 2022.
6. Mueller, Kristen L., Benjamin P. Cooper, Vicki Moran, Daphne Lew, Rachel Ancona, Joshua M. Landman, Marguerite Spruce, Phillip Marotta, David B. Liss, Michael A. Mancini, et al., "Incidence of and Factors Associated with Recurrent Firearm Injury Among Patients Presenting to St. Louis Trauma Centers, 2010 to 2019: A Cohort Study," *Annals of Internal Medicine*, Vol. 176, No. 9, 2023.
7. National Center for Health Statistics, "The National Death Index," webpage, July 18, 2024. As of July 19, 2024: https://www.cdc.gov/nchs/ndi/
8. National Center for Health Statistics, "NDI Costs and Payments," webpage, July 18, 2024: As of July 19, 2024: https://www.cdc.gov/nchs/ndi/apply/ndi-payments.html
9. National Center for Health Statistics, *National Death Index (NDI) Non-Economic Barriers Report: Strengthening the NDI For Research, Surveillance and Quality Assurance*, September 2019.
10. National Center for Health Statistics, "NDI Eligibility Requirements," webpage, July 18, 2024. As of July 19, 2024: https://www.cdc.gov/nchs/ndi/eligibility/index.html
11. National Center for Health Statistics, "NDI Application Review," webpage, July 18, 2024. As of July 19, 2024: https://www.cdc.gov/nchs/ndi/apply/review-process.html
12. Rothman, K. J., "The Epidemiologist's Lament," *American Journal of Public Health*, Vol. 71, No. 12, December 1981.
13. Hershey, N., "Putting the Lamentations of Epidemiologists in Perspective," *American Journal of Public Health*, Vol. 72, No. 10, October 1982.
14. Wartenberg, Daniel, and W. Douglas Thompson, "Privacy Versus Public Health: The Impact of Current Confidentiality Rules," *American Journal of Public Health*, Vol. 100, No. 3, March 2010.
15. U.S. Code, Title 42, Section 242m, General Provisions Respecting Effectiveness, Efficiency, and Quality of Health Services; paragraph (d), Information; Publication Restrictions.
16. U.S. Code, Title 44, Chapter 35, Coordination of Federal Information Policy; Sections 3561–3583.
17. National Center for Health Statistics, *National Death Index User's Guide*, 2024. As of August 21, 2024: https://www.cdc.gov/nchs/data/ndi/2024-NDI-User-Guide.pdf
18. Yan, Zhiyu, Victoria Nielsen, Glory Song, Anita Christie, Lee H. Schwamm, and Kori S. Zachrison, "Integration of Regional Hospitalizations, Registry and Vital Statistics Data for Development of a Single Statewide Ischemic Stroke Database," *Journal of Stroke and Cerebrovascular Diseases*, Vol. 31, No. 3, March 2022.
19. Valcarcel, Bryan, Sara J. Schonfeld, Christa L. Meyer, Ann Brunson, Julianne J. P. Cooley, Renata Abrahão, Ted Wun, Jeffery J. Auletta, Shahinaz M. Gadalla, Eric Engels, et al., "Comparison of Vital

Status, Cause of Death, and Follow-Up After Hematopoietic Cell Transplantation in Linked Center for International Blood and Marrow Transplant Research and California Cancer Registry Data, 1991 to 2018," *Transplantation and Cellular Therapy*, Vol. 30, No. 2, February 2024.

20. Loper, Ona R., Jessica M. Schultz, Tory L. King, Kimberly J. Tichy, Bethany F. Kintigh, and Donald J. Callaghan, "Benefits and Challenges of Sealing Deceased Individual Immunization Records in Immunization Information Systems (IIS): An Overview of Iowa's Process," *Journal of Public Health Management and Practice*, Vol. 30, No. 2, March–April 2024.
21. Heins, Sara, Zachary Predmore, Emily Hoch, and Sangita Baxi, *Linking Medicaid Claims, Birth Certificates, and Other Sources to Advance Maternal and Infant Health*, Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, October 2022.
22. Office of the Secretary Patient-Centered Outcome Research Trust Fund, *Enhancing Data Resources for Studying Patterns and Correlates of Mortality in PatientCentered Outcomes Research: Project 1 – Adding Cause-Specific Mortality to NCHS's National Hospital Care Survey by Linking to the National Death Index*, Project HP-17-002 (IAA #750117PE090019), February 2021.
23. Office of the Assistant Secretary for Planning and Evaluation, "Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research," webpage, undated. As of July 19, 2024: https://aspe.hhs.gov/enhancing-data-resources-researching-patterns-mortality-patient-centered-outcomes-research
24. Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, *Enhancing Data Resources for Studying Patterns and Correlates of Mortality in Patient-Centered Outcomes Research: Pilot Linkage of National Death Index+ to Commercially and Publicly insured Populations*, U.S. Food and Drug Administration, December 2020.