



Unlocking the Power of Mortality Data in Real-World Evidence: Overcoming Barriers and Driving Innovation

Phani Veeranki, MD, DrPH

Vice President, Optum Life Sciences

Jason LaBonte, PhD

CSO, Veritas Data Research

ISPOR Fast Fact Session

May 17, 2026; 11:30 AM- 12:00 PM



Conflict of Interest

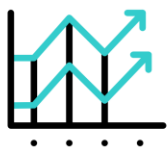
Phani Veeranki, MD, DrPH

- I have no conflicts of interest related to this presentation to declare.
- I am an employee of United Health Group (UHG).
 - The content is solely my responsibility and does not necessarily represent the official view of Optum/UHG or any other parties.

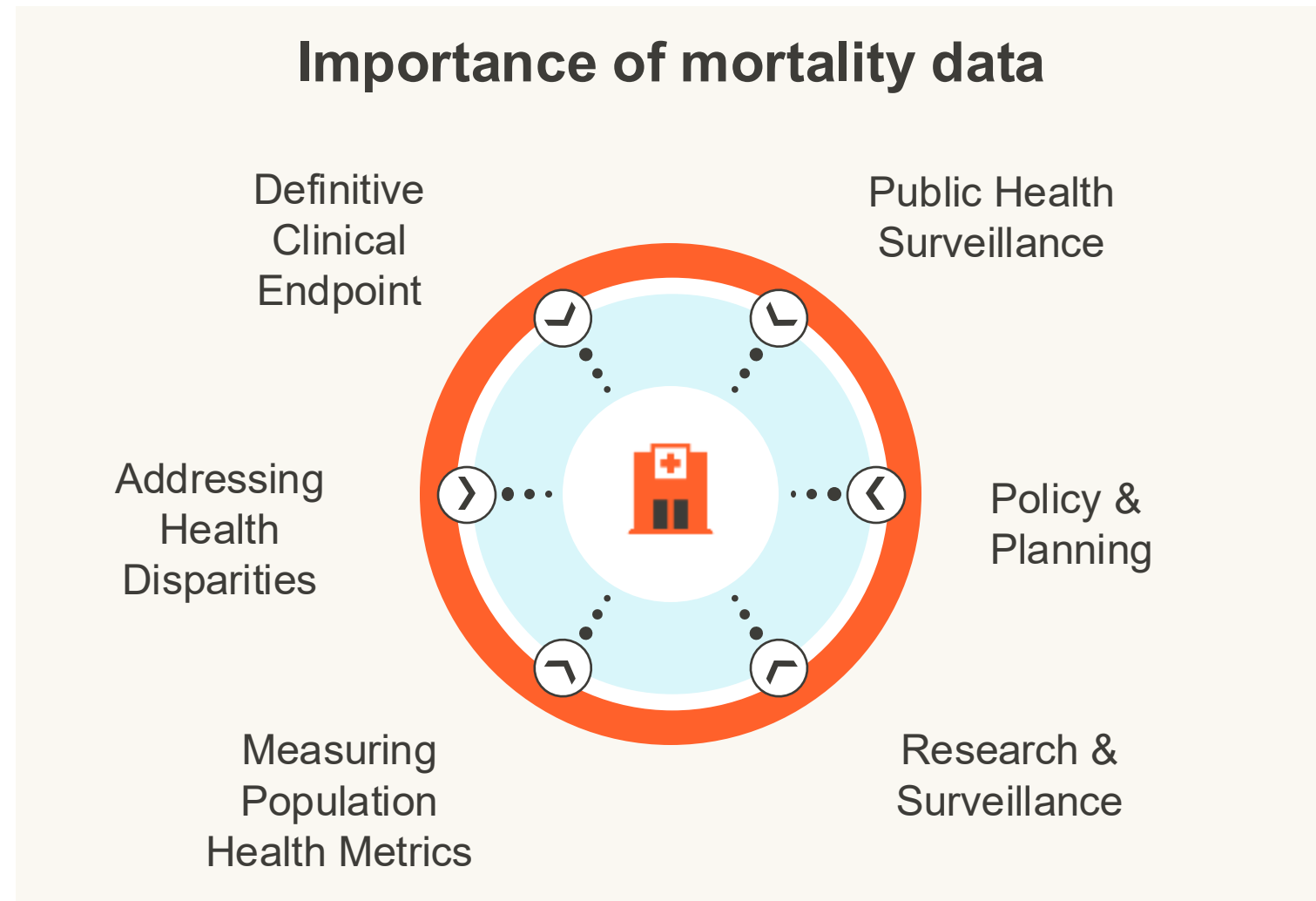
Jason LaBonte, PhD

- I have no conflicts of interest related to this presentation to declare.
- I am an employee of Veritas Data Research, Inc.

Mortality data and its importance



Mortality data are official statistics on the number (and causes) of deaths within a given population and timeframe.



Mortality data in Real-World Evidence



Mortality data isn't just a statistic—it's a powerful lens into real-world patient outcomes



Cornerstone for understanding the true impact of diseases and interventions in everyday clinical practice and diverse groups



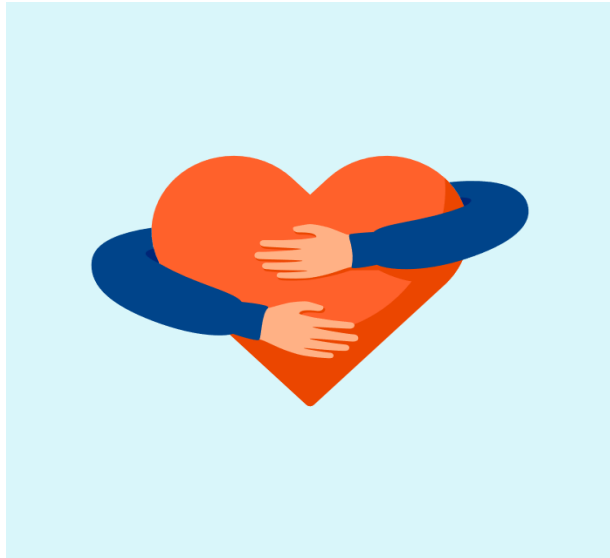
Unlike surrogate markers, mortality minimizes interpretational uncertainty in patient outcomes



Multiple data sources exist to track and verify death events

Mortality data in Real World Evidence

--*Utility*



Captures real-world effectiveness

Identifies mortality-related safety signals missed in trials

Reveals true U.S. disease burden through multi-source capture

Effectiveness

- Clinical trials often overestimate survival because they exclude older, frail, and comorbid patients and results are not representative of U.S. lung cancer populations.¹
- POSEIDON (Phase 3, 1L mNSCLC, Tremelimumab+Durvalumab+Chemo): 5-yr OS 15.7%.²

- The JCO Clinical Cancer Informatics 2023 study evaluated 3,882 melanoma patients.³
 - OS using EHR only = 17.63 years
 - OS using fully linked RWE (EHR + SSA DMF + obituaries + claims) = 13.39 years
- Median OS dropped from 17.63 → 13.39 years once missing deaths from SSA + obituaries + claims were added.

1. Ellis PM et al. A Narrative Review of the Strengths and Limitations of Real-World Evidence in Comparison to Randomized Clinical Trials: What Are the Opportunities in Thoracic Oncology for Real-World Evidence to Shine? *Curr Oncol.* 2025;32(11):629.
2. Peters S, Cho BC, Luft AV, et al. Durvalumab with or without tremelimumab + chemotherapy in first-line metastatic NSCLC (POSEIDON): 5-year OS outcomes. *J Thorac Oncol.* 2025;20(1):76-93.
3. Shao P, et al. Improving real-world mortality data quality in oncology research: augmenting electronic medical records with obituary, Social Security Death Index, and commercial claims data. *JCO Clin Cancer Inform.* 2023;7:e2300014.



“

RWE reveals real-world effectiveness is lower than suggested by trial OS.

”

Safety

- Trials Miss Late / Out-of-Facility Deaths

Once follow-up ends, deaths occurring after trial discontinuation are not routinely captured.

POSEIDON Phase 3 (Metastatic NSCLC) - After 63.4 months follow-up: No new safety signals were identified.¹

- Internet-Sourced Mortality (Obituaries, News, Public Notices) improved mortality detection sensitivity by 24% (Mass General Brigham) and 18% (Vanderbilt).

PPV for identifying deaths was 98.2%–98.9%, confirming accuracy.²

1. Peters S, Cho BC, Luft AV, et al. Durvalumab with or without tremelimumab + chemotherapy in first-line metastatic NSCLC (POSEIDON): 5-year OS outcomes. J Thorac Oncol. 2025;20(1):76-93.

2. LeNoue-Newton M et al. Augmenting fact and date of death in electronic health records using internet media sources: a validation study from two large healthcare systems. Am J Epidemiol. 2025 Nov 27.



“ RWE identifies additional real-world deaths that serve as safety signals (late toxicities, non-facility deaths, treatment-related declines). ”

Disease burden

- Trial populations are younger, healthier, and often lack comorbidities → underestimate true U.S. mortality burden.
- In FDA-submitted ICI trials for mNSCLC (13 trials, n = 9,285 US population), all enrolled patients were ECOG 0–1, meaning healthier, fitter, less comorbid populations.¹

- EHR captured 1,085 deaths, but obituary data added 1,127 more, and SSA DMF added 213 more.²
- Deaths identified across sources³:
 - Septicemia: 11.0% overlap, 61.1% claims-only, 27.9% SSA-only
 - Metastatic lung cancer: 7.5% overlap, 43.9% claims-only, 48.6% SSA-only
- RWE benefits by using multi-source linkage and provides true disease burden.

1. PeGoulart BHL, Mushti SL, Chatterjee S, et al. Correlations of RR & PFS with OS in metastatic NSCLC ICI trials (FDA pooled analysis; ECOG 0–1). *Lancet Oncol.* 2024;25(4):455-462.

2. Shao P, et al. Improving real-world mortality data quality in oncology research: augmenting electronic medical records with obituary, Social Security Death Index, and commercial claims data. *JCO Clin Cancer Inform.* 2023;7:e2300014.

3. Liu Y, Diakun D, Prinic N, Palmer L. Evaluation of mortality data sources for use in real-world data analyses (RWD154). *Value in Health.* 2024;27(S6):S387.



“ RWE uncovers hundreds of unreported deaths that no structured medical system recorded, revealing risks that RCT frameworks miss. ”

The background features several large, overlapping circles in shades of orange and light blue, creating a modern, abstract design. The text is centered in a dark grey, serif font.

Challenges with sourcing mortality data

Key Features of Mortality Data

Coverage

What percent of deaths can be found in the data set? If there are gaps, where are they (geographically, age groups, socioeconomic groups, ethnicities, etc.)?

Completeness

When data is available, which elements are included (names, dates of birth, SSN or other identifying attributes, location of death, cause of death (COD), etc.)?

Timeliness

How fast is a death record available after a mortality event?

Transparency / Traceability

What is the reliability of the death record? Can it be traced back to the originating source to prove its validity? How often do elements of the death record change?

Accessibility

Are you allowed to use the data for your purpose, and what format does it come in?

Sources of Mortality Data

Mortality data can be gathered from multiple sources

Each source has its own characteristics

Government reporting:

- Individual states
- CDC's National Death Index
- Social Security Administration's Death Master File

Extraction from medical data:

- EHR data
- Medical claims data

Public announcements:

- Obituaries
- Interment directories

State Vital Statistics

PROs

- Death certificates are collected by each state, making them an authoritative source
- Generally, no bias due to the completeness of the data and traceability is high

CONs

- Each state has different access rules for its data:
 - Some states only allow access for IRB-approved projects (e.g. Florida)
 - Some states only report partial data (e.g. Connecticut provides COD, most don't)
 - Some states only allow certain use cases, like fraud prevention
 - Some states prohibit any data sharing (e.g. Colorado)
- To access data, you need to apply to each vital stats office individually. Distribution of mortality data is subject to state regulations (generally dictated by the state legislature, meaning it is very slow to change)
- Cause of death accuracy has been shown to be lacking (J. Alipour, 2021)

CDC's National Death Index (NDI)

PROs

- States send vital statistics data to CDC, making it the most complete and unbiased mortality data set
- CDC makes aggregate mortality statistics available publicly (<https://wonder.cdc.gov/>)

CONs

- NDI services are available only to researchers for statistical purposes in public health and medical studies
- NDI fees vary based on the number of study subjects, the type of search requested, and the length of time to be searched
- Timeliness is slow, and CDC performs all matching to NDI, which adds time as well

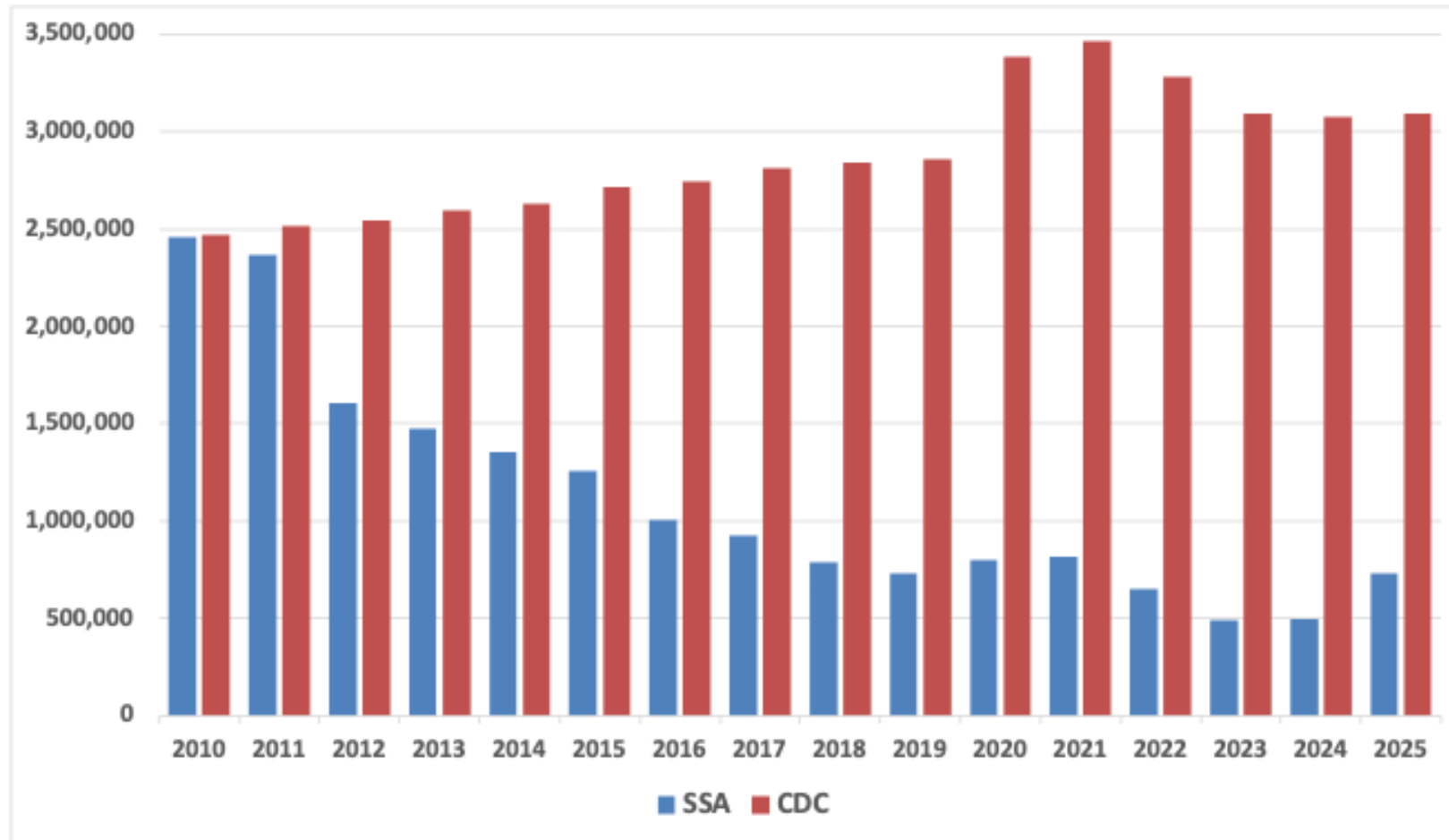
Social Security Administration's Death Master File

PROs

- Historical data is very complete

CONs

- Legislation in 2011 removed state records, causing substantial drop in coverage
- DOGE efforts marked 6 million records with default dates through 2025, causing implausible ages at death (often 120+ years old)
- Accessibility can be low, with a lengthy application



Deaths Reported in EHRs

PROs

- Mortality information is typically recorded in an EHR when a patient dies during in-patient care, including hospice, making it a high confidence source
- Deaths can be directly coupled to medical data, allowing derivation of cause of death

CONS

- Not every patient dies in the hospital. UCLA performed a study on their EHR showing that 20% their seriously diseased patients were dead and they didn't know it
- While individual EHRs might have good coverage, aggregate EHR data from every health system isn't available. One aggregated source had only 6-8% of national deaths
- EHRs are PHI, so deaths extracted from EHRs are typically only available in deidentified format
- Deidentified EHR records are difficult to verify for traceability

Deaths Extracted from Medical Claims

PROs

- Deaths may be reported in medical claims as a discharge codes for expiry and occurrence codes for the date of death
- Aggregated claims data is much more available, and can be nationally representative
- Claims are updated daily in some cases, making it very timely data

CONs

- Not every patient dies in the hospital, and not every death has a claim filed
- One aggregated claims source had 18-24% of national deaths
- Claims are PHI, so deaths extracted from claims are typically only available in deidentified format
- Deidentified claims are difficult to verify for traceability

Public Announcements of Deaths

PROs

- Deaths are reported publicly in obituaries, funeral notices, memorials, graveyard directories, and more
- Data is free, generally timely (e.g. within days of a death for obituaries), and has no restrictions on usage
- Data traceability is high, as you can point back to the originating source

CONs

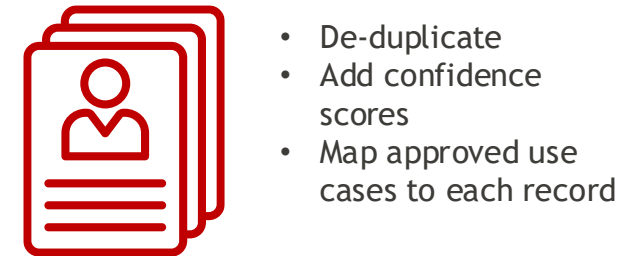
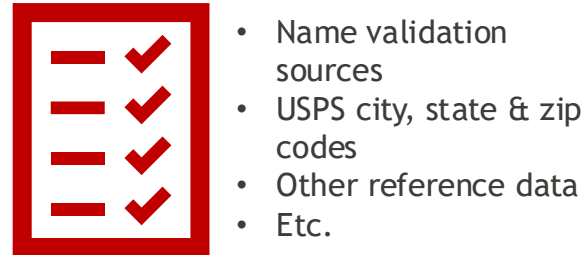
- Data is very fragmented (reported across ~40,000 different websites and newspapers)
- Data formats and completeness vary
- Data is often self-reported, so accuracy can vary

Comparison of Mortality Sources

	National Coverage	Complete	Timeliness	Transparent	Accessible
Individual States					
CDC					
SSA					
EHR					
Claims					
Individual Public Sources					

Veritas's Approach: Combining Multiple Sources

An example of how to combine 40,000 sources into one daily-updated dataset



SCAN

Ingest mortality data from online and offline sources daily.

CLEAN

Using third-party reference data, enhance individual records

CURATE

Validate records across sources and merge using proprietary rules