

Daten des Übersterblichkeitsberichts

Robert Koch-Institut¹

Beitragende

Benedikt Zacher², Andreas Hicketier², Alexander Ullrich², Michaela Diercke², & Matthias an der Heiden³

¹ Bundesministerium für Gesundheit (BMG)

² Robert Koch-Institut | Fachgebiet 32

³ Robert Koch-Institut | Fachgebiet 34

Zitieren

Robert Koch-Institut. (2025). Daten des Übersterblichkeitsberichts [Data set]. Zenodo.

<https://doi.org/10.5281/zenodo.17463782>

Zusammenfassung

Im Datensatz "Daten des Übersterblichkeitsberichts" des Robert Koch-Instituts werden wöchentliche Daten zur Übersterblichkeit in Deutschland bereitgestellt. Der Datensatz enthält Angaben zum Wochendatum und zur jeweiligen Gruppe (z. B. Altersgruppen, Geschlecht oder Region). Für jede Gruppe wird die beobachtete Sterbefallzahl den modellbasiert erwarteten Sterbefallzahlen gegenübergestellt. Zur Bewertung von Abweichungen zwischen beobachteter und erwarteter Sterbefallzahl wird der z-Score berechnet, der zur Einordnung einer etwaigen Übersterblichkeit genutzt wird. Der Datensatz dient der kontinuierlichen Überwachung von Sterbefallzahlen in Deutschland und ermöglicht die frühzeitige Erkennung von ungewöhnlich hohen Sterbefallzahlen in den betrachteten Gruppen.

Inhaltsverzeichnis

- [Informationen zum Datensatz und Entstehungskontext](#)
- [Aufbau und Inhalt des publizierten Datensatzes](#)
- [Hinweise zur Nachnutzung der Daten](#)

Informationen zum Datensatz und Entstehungskontext

Ereignisse wie Infektionswellen oder Hitzeperioden führen regelmäßig zu kurzfristigen Anstiegen der Sterbefallzahlen in Deutschland. Wenn die Zahl der Sterbefälle im Vergleich zur erwarteten Sterbefallzahl deutlich erhöht ist, wird von Übersterblichkeit gesprochen.

Im Rahmen der Mortalitätssurveillance werden nach § 5b BevStatG dieser Datensatz sowie ein Bericht zur Einordnung veröffentlicht. Wie ermöglichen eine zeitnahe Einschätzung der aktuellen Sterbefallzahlen in Deutschland. Der Datensatz wird wöchentlich aktualisiert, um fortlaufend über die aktuelle Lage zur Sterbefallzahlen und Übersterblichkeit zu informieren.

Der Bericht zur Übersterblichkeit in Deutschland ist verfügbar unter:

Robert Koch-Institut. (2025). Bericht zur Übersterblichkeit in Deutschland. Zenodo.
<https://doi.org/10.5281/zenodo.17608501>

Administrative und organisatorische Angaben

Die zugrundeliegenden Daten für die Mortalitätssurveillance gemäß § 5b Bevölkerungsstatistikgesetz (BevStatG) werden vom [Statistischen Bundesamt \(Destatis\)](#) an das [Robert Koch-Institut \(RKI\)](#) übermittelt. Die Konzeption und Durchführung der Analysen erfolgen im [Fachgebiet 32 | Surveillance und elektronisches Melde- und Informationssystem \(DEMIS\) | ÖGD-Kontaktstelle](#) des RKI.

Die Veröffentlichung der aufbereiteten und validierten Daten, die Kuration der Datensätze sowie das Qualitätsmanagement der Metadaten werden durch das Fachgebiet [MF 4 | Fach- und Forschungsdatenmanagement](#) des RKI verantwortet. Fragen zum Datenmanagement oder zu den offenen Datensätzen können an das Open-Data-Team des Fachgebiets MF 4 unter OpenData@rki.de gerichtet werden.

Datenerhebung und Datenauswertung

Gemäß § 5b des BevStatG übermittelt Destatis dem RKI für epidemiologische Analysen einer überdurchschnittlichen Sterblichkeit Einzelangaben zu jedem registrierten Sterbefall. Diese Angaben umfassen das Geschlecht, Jahr und Monat der Geburt, den Sterbetag, die Anschrift des Sterbeortes, den Landkreis bzw. die kreisfreie Stadt der letzten Meldung sowie das zuständige Standesamt, das den Sterbefall registriert hat. Die Datengrundlage dieses Übersterblichkeitsberichts entspricht derjenigen des Sterbefallmonitorings des Statistischen Bundesamts.

Es werden zwei statistische Verfahren für die Analyse angewendet: die Hochrechnung der Fallzahlen (Nowcasting) und die Übersterblichkeitsanalyse.

Da Sterbefallmeldungen zeitverzögert eingehen, wird ein Schätzverfahren eingesetzt, um die endgültigen Fallzahlen einer Sterbewoche frühzeitig zu prognostizieren. Hierzu wird ein generalisiertes lineares quasi-Poisson-Modell verwendet, das den Zusammenhang zwischen bereits gemeldeten und endgültigen Fallzahlen beschreibt. Das Modell berücksichtigt Zeittrends im Meldeverhalten sowie Feiertagseffekte und wird auf historischen Daten trainiert.

Zur Identifikation von Phasen erhöhter Sterblichkeit werden Hidden Markov Modelle eingesetzt, die zwischen einem normalen Sterblichkeitsniveau und verschiedenen Zuständen erhöhter Sterblichkeit unterscheiden. Das Modell beinhaltet ein generalisiertes linearen Modell, das die erwartete Sterbefallzahl unter Berücksichtigung von Saisonalität und Zeittrends schätzt. Anschließend ein z-Score berechnet, analog zur Methodik von [EuroMOMO](#), wodurch internationale Vergleichbarkeit gewährleistet ist.

Die zugrundeliegende Methodik ist als Open Source R-Paket [excode auf github](#) und Zenodo verfügbar.

Aufbau und Inhalt des publizierten Datensatzes

Der Datensatz enthält wöchentliche Informationen zur Sterblichkeit in Deutschland. Jede Zeile beschreibt eine Kombination aus Wochendatum und Gruppe (z. B. Altersgruppe, Geschlecht oder Region). Es werden die beobachtete Sterbefallzahl, die erwartete Sterbefallzahl sowie der daraus abgeleitete z-Score angegeben. Ergänzend wird angegeben, ob es sich um hochgerechnete Werte handelt.

Variablen und Variablenausprägungen

Die Datei [daten_uebersterblichkeitsbericht.tsv](#) enthält die in der folgenden Tabelle abgebildeten Variablen und deren Ausprägungen. Ein maschinenlesbares Datenschema ist im [Data Package Standard](#) in [tableschema_daten_uebersterblichkeitsbericht.json](#) hinterlegt: > [tableschema_daten_uebersterblichkeitsbericht.json](#)

| Variable | Typ | Ausprägungen | Beschreibung |
|--------------------------|---------|--|--|
| Wochendatum | date | Format: YYYY-MM-DD Fehlende Werte: NA | Datum (Sonntag) der Sterbewoche |
| Gruppe | string | Werte: 0-14 Jahre, 15-44 Jahre, 45-64 Jahre, 65-74 Jahre, 75-84 Jahre, 85+ Jahre, Alten- und Pflegeheime, ... Beispiel: 45-64 Jahre Fehlende Werte: NA | Stratifizierung nach Region (Deutschland, Bundesländer), Altersgruppen und Sterbeort (Krankenhäuser, Alten- und Pflegeheime und andere Adressen) |
| Sterbefallzahl | integer | Werte: ≥0 Beispiel: 2032 Fehlende Werte: NA | Anzahl der Verstorbenen in der jeweiligen Woche und Gruppe |
| Erwartete_Sterbefallzahl | number | Werte: ≥0 Beispiel: 2109.17949033554 Fehlende Werte: NA | Erwartete Anzahl der Verstorbenen in der jeweiligen Woche und Gruppe |
| zscode | number | Beispiel: -1.23644832483487 Fehlende Werte: NA | Der z-score gibt an, wie weit ein Wert vom Durchschnitt entfernt ist, gemessen in Standardabweichungen. |

| Variable | Typ | Ausprägungen | Beschreibung |
|--------------|--------|--|---|
| Hochrechnung | string | Werte: ja , nein Beispiel: nein Fehlende Werte: NA | "ja" falls die Sterbefallzahl hochgerechnet, also für den Meldeverzug korrigiert, wurde. "nein" andernfalls. |

Formatierung

Die Daten sind im Datensatz als Tab-separierte .tsv-Datei enthalten. Der verwendete Zeichensatz ist UTF-8.

- Zeichensatz: UTF-8
- .tsv-Trennzeichen: Tab "\t"

Metadaten

Zur Erhöhung der Auffindbarkeit sind die bereitgestellten Daten mit Metadaten beschrieben. Über GitHub Actions werden Metadaten an die entsprechenden Plattformen verteilt. Für jede Plattform existiert eine spezifische Metadatendatei, diese sind im Metadatenordner hinterlegt:

Metadaten/

Versionierung und DOI-Vergabe erfolgt über [Zenodo.org](#). Die für den Import in Zenodo bereitgestellten Metadaten sind in der [zenodo.json](#) hinterlegt. Die Dokumentation der einzelnen Metadatenvariablen ist unter <https://developers.zenodo.org/#representation> nachlesbar.

Metadaten/zenodo.json

In der zenodo.json ist neben dem Publikationsdatum ("publication_date") auch der Datenstand in folgendem Format enthalten (Beispiel):

```

"dates": [
  {
    "start": "2023-09-11T15:00:21+02:00",
    "end": "2023-09-11T15:00:21+02:00",
    "type": "Created",
    "description": "Date when the published data was created"
  }
],
  
```

Zusätzlich beschreiben wir tabellarische Daten mithilfe des [Data Package Standards](#). Ein Data Package ist eine strukturierte Sammlung von Daten und zugehörigen Metadaten, die den Austausch und die Wiederverwendung von Daten erleichtert. Es besteht aus einer datapackage.json-Datei, die zentrale Informationen wie die enthaltenen Ressourcen, ihre Formate und Schema-Definitionen beschreibt.

Der Data Package Standard wird von der [Open Knowledge Foundation](#) bereitgestellt und ist ein offenes Format, das eine einfache, maschinenlesbare Beschreibung von Datensätzen ermöglicht.

Die Liste der in diesem Repository enthaltenen Daten ist in folgender Datei hinterlegt:

datapackage.json

Für tabellarische Daten definieren wir zusätzlich ein [Table Schema](#), das die Struktur der Tabellen beschreibt, einschließlich Spaltennamen, Datentypen und Validierungsregeln. Diese Schema-Dateien finden sich unter:

Metadaten/schemas/

Hinweise zur Nachnutzung der Daten

Offene Forschungsdaten des RKI werden auf [Zenodo.org](#), [GitHub.com](#), [OpenCoDE](#) und [Edoc.rki.de](#) bereitgestellt:

- <https://zenodo.org/communities/robertkochinstitut>
- <https://github.com/robert-koch-institut>
- <https://gitlab.opencode.de/robert-koch-institut>
- <https://edoc.rki.de/>

Lizenz

Der Datensatz "Daten des Übersterblichkeitsberichts" ist lizenziert unter der [Creative Commons Namensnennung 4.0 International Public License | CC-BY 4.0 International](#).

Die im Datensatz bereitgestellten Daten sind, unter Bedingung der Namensnennung des Robert Koch-Instituts als Quelle, frei verfügbar. Das bedeutet, jede Person hat das Recht die Daten zu verarbeiten und zu verändern, Derivate des Datensatzes zu erstellen und sie für kommerzielle und nicht kommerzielle Zwecke zu nutzen. Weitere Informationen zur Lizenz finden sich in der [LICENSE](#) bzw. [LIZENZ](#) Datei des Datensatzes.