

SARI-Hospitalisierungsinzidenz

Kristin Tolksdorf¹, Luise Goerlitz¹, Doreen Staat¹, Tamar Gvaladze¹, Walter Haas¹, & Silke Buda¹

¹ Robert Koch-Institut | Fachgebiet 36

Zitieren

Tolksdorf, K., Goerlitz, L., Staat, D., Gvaladze, T., Haas, W., & Buda, S. (2025). SARI-Hospitalisierungsinzidenz [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.17540177>

Zusammenfassung

Im Datensatz "SARI-Hospitalisierungsinzidenz" des Robert Koch-Instituts werden die wöchentlichen Inzidenzen schwerer akuter respiratorischer Infektionen (SARI) auf Bundesebene in Deutschland bereitgestellt. Die Daten stammen aus der syndromischen Krankenhaussurveillance ICOSARI, die auf ICD-10-Diagnosecodes aus etwa 70 Sentinelkliniken basiert. Der Datensatz umfasst Hospitalisierungsinzidenzen pro 100.000 Einwohner, aufgeschlüsselt nach Altersgruppen, Saison und Kalenderwoche. Neben der Gesamtheit aller SARI-Fälle werden auch Fälle mit zusätzlicher COVID-19-, Influenza- oder RSV-Diagnose separat erfasst. Die Surveillance liefert entscheidende Informationen zur Krankheitslast akuter respiratorischer Erkrankungen und dient als Grundlage für epidemiologische Bewertungen und Empfehlungen für gesundheitspolitische Maßnahmen.

Inhaltsverzeichnis

- Informationen zum Datensatz und Entstehungskontext
- Aufbau und Inhalt des Datensatzes
- Hinweise zur Nachnutzung der Daten

Informationen zum Datensatz und Entstehungskontext

Zur Einschätzung der Krankheitslast schwerer symptomatischer Atemwegsinfektionen im stationären Bereich wird mithilfe von Daten aus der syndromischen Krankenhaussurveillance ICOSARI die Inzidenz der Fälle, die mit einer schweren akuten respiratorischen Infektion (SARI) in ein Krankenhaus zur Behandlung aufgenommen wurden, pro 100.000 Einwohner berechnet.

Zudem werden SARI-Fälle pro 100.000 Einwohner, bei denen zusätzlich zur SARI eine Erkrankung mit COVID-19, mit Influenza oder mit RSV diagnostiziert wurde, separat dargestellt. Zeitnahe und valide Daten über die Häufigkeit von SARI in der Bevölkerung sowie zum Auftreten von COVID-19-, Influenza- und RSV-Erkrankungen unter SARI-Patientinnen und -Patienten im zeitlichen Verlauf sind essenziell für die Einschätzung der epidemiologischen Lage und die Entwicklung und Anpassung von Präventionsstrategien.

Administrative und organisatorische Angaben

Die zugrundeliegenden Daten werden im Rahmen einer wissenschaftlichen Kooperation vom Datenzentrum der HELIOS Kliniken GmbH dem Robert Koch-Institut (RKI) zur Verfügung gestellt. Die Konzeptionierung der Datennutzung, das Datenmanagement, die Validierung und Analyse der Daten im Rahmen der SARI-Surveillance sowie die fachliche Bewertung der Ergebnisse erfolgen im [Fachgebiet 36 | Respiratorisch übertragbare Erkrankungen](#) des RKI.

Die Veröffentlichung der Analyseergebnisse, die Kuration sowie das Qualitätsmanagement der (Meta-)Daten erfolgt durch das [Fachgebiet MF 4 | Fach- und Forschungsdatenmanagement](#). Fragen zum Meta-Datenmanagement können an das Open Data-Team des Fachgebiets MF4 gerichtet werden (OpenData@rki.de).

Entstehungskontext

Die ganzjährige syndromische Surveillance schwerer akuter respiratorischer Infektionen (SARI) informiert über die aktuelle Krankheitsschwere und -häufigkeit.

In Deutschland erfolgt die syndromische Surveillance schwerer akuter respiratorischer Erkrankungen im stationären Bereich durch die syndromische Krankenhaus surveillance ICOSARI (ICD-10-Code basierte Krankenhaus surveillance schwerer akuter respiratorischer Infektionen) RKI. Diese basiert auf fallbasierten anonymisierten Datensätzen zu ärztlichen Diagnosen aus Sentinel-Krankenhäusern.

Die Ergebnisse tragen zum Verständnis der Epidemiologie der Influenza, COVID-19, RSV-Infektionen und von weiteren akuten respiratorischen Erkrankungen sowie zur Entwicklung von Präventionsstrategien bei. Eine kontinuierliche SARI-Surveillance im stationären Bereich wird vom [ECDC](#) und der [WHO](#) empfohlen.

Datenerhebung

Derzeit werden Daten aus aktuell ca. 70 Sentinelkliniken mit einer bundesweiten Abdeckung von 5-6 % aller in Deutschland hospitalisierten Patientinnen und Patienten erhoben. In den teilnehmenden Krankenhäusern werden fallbezogene Daten über aufgenommene Patientinnen und Patienten routinemäßig erfasst. Die Daten aus diesen Krankenhäusern werden zentral im Datenzentrum im Rahmen des Qualitätsmanagements validiert. Von dort aus werden validierte, anonymisierte, fallbezogene Datensätze von neu im Krankenhaus aufgenommenen Patientinnen und Patienten an das RKI übermittelt.

Buda, S., Tolksdorf, K., Schuler, E. et al. Establishing an ICD-10 code based SARI-surveillance in Germany – description of the system and first results from five recent influenza seasons. BMC Public Health 17, 612 (2017). DOI: [10.1186/s12889-017-4515-1](https://doi.org/10.1186/s12889-017-4515-1)

Der hier zur Verfügung gestellte Datensatz enthält die Werte der wöchentlichen Hospitalisierungsinzidenz von Patientinnen und Patienten in sechs Altersgruppen mit den ICD-10-Diagnosen J09 bis J22. Das sind Diagnose-Codes für akute Infektionen der unteren Atemwege. Zusätzlich sind SARI-Fälle, die einen zusätzlichen krankheitsspezifischen Diagnosecode für eine COVID-19, Influenza- oder RSV-Erkrankung erhalten haben, aufgeführt.

Datenauswertung und -aufbereitung

Zur Einschätzung der Krankheitslast schwerer symptomatischer Erkrankungen im stationären Bereich wird mithilfe der Daten aus dem ICOSARI-Sentinel wöchentlich die Inzidenz der Fälle berechnet, die in ein Krankenhaus aufgenommen wurden und dort die Diagnose einer schweren akuten Atemwegserkrankung erhalten haben (SARI-Hospitalisierungsinzidenz).

Hierbei werden Patientinnen und Patienten pro Kalenderwoche der Krankenhausaufnahme erfasst, die einen ICD-10-Code für SARI ([ICD-10-Codes J09 bis J22](#)) in der Hauptdiagnose erhalten.

Zudem werden drei Untergruppen der SARI-Hospitalisierungsinzidenz (Gesamt) dargestellt:

- Inzidenz von SARI-Fällen mit einem ICD-10-Diagnosecode für eine **COVID-19-Erkrankung** (U07.1) in einer der Nebendiagnosen (COVID-SARI)
- Inzidenz von SARI-Fällen mit einem ICD-10-Diagnosecode für eine **Influenza-Erkrankung** (J10) in der Haupt- oder einer der Nebendiagnosen (Influenza-SARI)
- Inzidenz von SARI-Fällen mit einem ICD-10-Diagnosecode für eine **RSV-Erkrankung** (J12.1, J20.5, J21.0) in der Haupt- oder einer der Nebendiagnosen (RSV-SARI)

Die Berechnung der wöchentlichen SARI-Hospitalisierungsinzidenz erfolgte wie von [Tolksdorf et al \(2022\)](#) beschrieben. Detaillierte Informationen zu Datenerhebung und Auswertung lassen sich folgenden Publikationen entnehmen:

Cai W, Kondgen S, Tolksdorf K, Dürrwald R, Schuler E, Biere B, Schweiger B, Goerlitz L, Haas W, Wolff T, Buda S, Reiche J. Atypical age distribution and high disease severity in children with RSV infections during two irregular epidemic seasons throughout the COVID-19 pandemic, Germany, 2021 to 2023. Euro Surveill. 2024 Mar;29(13):2300465. [DOI: 10.2807/1560-7917.ES.2024.29.13.2300465](#)

Buchholz U, Lehfeld AS, Tolksdorf K, Cai W, Reiche J, Biere B, Dürrwald R, Buda S. Respiratory infections in children and adolescents in Germany during the COVID-19 pandemic. J Health Monit. 2023 Jun 14;8(2):20-38. [DOI: 10.25646/11437](#)

Tolksdorf K, Haas W, Schuler E, Wieler LH, Schilling J, Hamouda O, Diercke M, Buda S (2022): Syndromic surveillance for severe acute respiratory infections (SARI) enables valid estimation of COVID-19 hospitalization incidence and reveals underreporting of hospitalizations during pandemic peaks of three COVID-19 waves in Germany, 2020-2021 [DOI: 10.1101/2022.02.11.22269594](#)

Tolksdorf K, Buda S, Schuler E, Wieler LH, Haas W (2020): Eine höhere Letalität und lange Beatmungsdauer unterscheiden COVID-19 von schwer verlaufenden Atemwegsinfektionen in Grippewellen. Epid Bull 2020;41:3-10 | [DOI: 10.25646/7111](#)

Cai W, Tolksdorf K, Hirve S, Schuler E, Zhang W, Haas W, Buda S. Evaluation of using ICD-10 code data for respiratory syncytial virus surveillance. Influenza Other Respi Viruses.2020;14:630-637. [DOI: 10.1111/irv.12665](#)

Goerlitz L, Tolksdorf K, Buchholz U, Prahm K, Preuß U, An der Heiden M, Wolff T, Dürrwald R, Nitsche A, Michel J, Haas W, Buda S. Überwachung von COVID-19 durch Erweiterung der etablierten Surveillance für Atemwegsinfektionen [Monitoring of COVID-19 by extending existing surveillance for acute respiratory infections]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2021 Apr;64(4):395-402. German. [DOI: 10.1007/s00103-021-03303-2](#)

Buda S, Tolksdorf K, Schuler E, Kuhlen R, Haas W. (2017): Establishing an ICD-10 code based SARI-surveillance in Germany – description of the system and first results from five recent influenza seasons. BMC Public Health 17, 612 (2017). [DOI: 10.1186/s12889-017-4515-1](#)

Limitationen und Einordnung des Datensatzes

Die Daten haben zwar eine eingeschränkte geografische Auflösung, sie erlauben jedoch bundesweit robuste Aussagen zur Krankheitslast schwerer akuter Atemwegserkrankungen mit COVID-19, Influenza- oder RSV- Erkrankung für unterschiedliche Altersgruppen.

Aufbau und Inhalt des Datensatzes

Der Datensatz enthält die wöchentlichen Werte der SARI-Hospitalisierungsinzidenz, die im Rahmen der syndromischen SARI-Surveillance im ICOSARI-Sentinel des RKI berechnet wurden. Im Datensatz enthalten sind:

- wöchentliche Hospitalisierungsinzidenz von schweren akuten respiratorischen Erkrankungen pro 100.000 Einwohner in Deutschland (Gesamt, COVID-19, Influenza, RSV)
- Lizenz-Datei mit der Nutzungs Lizenz des Datensatzes in Deutsch und Englisch
- Datensatzdokumentation in deutscher Sprache
- Metadaten zur automatisierten Weiterverarbeitung

Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen auf Bundesebene

Die Daten der Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen sind nach folgenden Merkmalen differenziert:

- Saison
- Kalenderwoche
- Altersgruppe
- SARI (Gesamt, COVID-19, Influenza, RSV)

Die Daten werden dienstags im Rahmen der wöchentlichen Berichterstattung ausgewertet. Das bedeutet, dass alle bis dahin am RKI validierten Daten einfließen. Die Daten sind bundesweit wöchentlich verfügbar und können durch Nachmeldungen noch ergänzt werden.

SARI-Hospitalisierungsinzidenz.tsv

Variablen und Variablenausprägungen

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

tableschema_SARI-Hospitalisierungsinzidenz.json

Variable	Typ	Ausprägungen	Beschreibung
Saison	date	Format: YYYY/YY	Saison jeweils von Kalenderwoche 40 bis Kalenderwoche 39 des Folgejahres (z.B. Saison 2012/13 = 2012-W40 bis 2013-W39)
Kalenderwoche	date	Format: YYYY-Www	Kalenderwoche der Krankenhausneuaufnahme im ISO-8601 Format
Altersgruppe	string	Werte: 00+ , 0-4 , 5-14 , 15-34 , 35-59 , 60-79 , 80+	Altersgruppen in Jahresintervallen. Zusätzlich entsprechen die Gruppen 00+ : Personen jeden Alters und 80+ : Personen ab 80 Jahren

SARI	string	Werte: Gesamt , COVID-19, Influenza , RSV	Erkrankung, auf welche sich die Hospitalisierungsinzidenz bezieht. Gesamt : Alle erhobenen SARI-Erkrankungen (einschließlich COVID-19, Influenza, RSV und andere)
Hospitalisierungsinzidenz	number	Werte: ≥ 0	Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen (SARI) je 100.000 Einwohner

Formatierung

Die Daten sind im Datensatz als tabseparierte Datei (TSV) enthalten. Der verwendete Zeichensatz der TSV-Datei ist UTF-8. Trennzeichen der einzelnen Werte ist ein Tab "\t".

- 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": "Collected",
    "description": "Date when the Dataset 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 "SARI-Hospitalisierungsinzidenz" 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.