

COVID-SARI-Hospitalisierungsinzidenz

Robert Koch-Institut | RKI

Nordufer 20

13353 Berlin

Kristin Tolksdorf¹, Luise Goerlitz¹, Tamar Gvaladze¹, Walter Haas¹ und Silke Buda

¹ Robert Koch-Institut | Fachgebiet 36

Zitieren

Tolksdorf K, Goerlitz L, Gvaladze T, Haas W und Buda S (2024): COVID-SARI-Hospitalisierungsinzidenz, Berlin:Zenodo. DOI:[10.5281/zenodo.13943810](https://doi.org/10.5281/zenodo.13943810)

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) und einer COVID-19-Diagnose in ein Krankenhaus zur Behandlung aufgenommen wurden, pro 100.000 Einwohner berechnet. Zeitnahe und valide Daten über die Häufigkeit von SARI mit COVID-19 im zeitlichen Verlauf sind essenziell für die Einschätzung der epidemiologischen Lage und die Entwicklung und Anpassung von Präventionsstrategien auch nach dem Ende der COVID-19-Pandemie.

Administrative und organisatorische Angaben

Die zugrunde liegenden 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 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 Krankheitslast und -häufigkeit. In Deutschland erfolgt die syndromische Surveillance schwerer akuter respiratorischer Erkrankungen im stationären Bereich durch die syndromische Krankenhaussurveillance ICOSARI (ICD-10-Code basierte Krankenhaussurveillance schwerer akuter respiratorischer Infektionen) des Robert Koch-Instituts (RKI). Diese basiert auf fallbasierten anonymisierten DRG-Datensätzen 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 hospitalisierter Patient:innen 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 anonymisierte validierte, anonymisierte, fallbezogene Datensätze von neu im Krankenhaus aufgenommenen Patient:innen und Patienten an das Robert Koch-Institut RKI übermittelt ([Establishing an ICD-10 code based SARI-surveillance in Germany - description of the system and first results from five recent influenza seasons](#)).

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 sowie eine Diagnose von COVID-19 erhalten haben (COVID-SARI-Hospitalisierungsinzidenz).

COVID-SARI-Falldefinition

Der hier zur Verfügung gestellte Datensatz enthält die Werte der wöchentlichen Hospitalisierungsinzidenz von Patientinnen und Patienten mit einem ICD-10-Diagnosecode für eine SARI (J09 bis J22, Diagnose-Codes für akute Infektionen der unteren Atemwege) sowie einem ICD-10-Diagnosecode für eine COVID-19-Erkrankung (U07.1).

- SARI-Diagnose als Hauptdiagnose UND eine COVID-19 Diagnose

Änderung der Falldefinition zum 12.10.2023

Zum 12.10.2023 wurde die Falldefinition der COVID-SARI-Hospitalisierungsinzidenz geändert. Die Datenreihen gemäß der vorherigen Falldefinition (SARI in der Haupt- oder einer der Nebendiagnosen UND Diagnose einer laborbestätigten COVID-19 Erkrankung) stehen weiterhin zum Download zur Verfügung (DOI: [10.5281/zenodo.8435968](#)), werden aber seit dem 12.10.2023 nicht fortgeführt. Weitere Informationen finden sich in den [Hinweisen zur Anpassung der Falldefinition](#).

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:

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](#)

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: [10.1007/2Fs00103-021-03303-2](https://doi.org/10.1007/2Fs00103-021-03303-2).

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](https://doi.org/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.

Aufbau und Inhalt des Datensatzes

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

- wöchentliche Hospitalisierungsinzidenz von schweren akuten respiratorischen Erkrankungen mit zusätzlicher COVID-19-Diagnose pro 100.000 Einwohner in Deutschland
- Lizenz-Datei mit der Nutzungslizenz des Datensatzes in Deutsch und Englisch
- Datensatzdokumentation in deutscher Sprache
- Metadaten zur automatisierten Weiterverarbeitung

Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen mit COVID-19 auf Bundesebene

Die Daten der Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen mit COVID-19 sind nach folgenden Merkmalen differenziert:

- Kalenderwoche

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.

[COVID-SARI-Hospitalisierungsinzidenz.csv](#)

Variablen und Variablenausprägungen

Variable	Typ	Ausprägung	Beschreibung
date	Datum	jjjj-Www	Kalenderwoche der Krankenhausneuaufnahme im ISO-8601 Format
agegroup	Text	60-79 , 80+ , 00+	Altersgruppen in Jahren, 00+ gibt die Gesamtinzidenz über alle Altersgruppen an.
sari_covid19_incidence	Rationale Zahl	>0	Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen mit COVID-19

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 Metadaten-datei, diese sind im Metadatenordner hinterlegt:

Metadaten/

Versionierung und DOI-Vergabe erfolgt über [Zenodo.org](https://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 der Publikationsdatum (`"publication_date"`) auch der Datenstand enthalten:

```
"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"  
  }  
],
```

Hinweise zur Nachnutzung der Daten

Offene Forschungsdaten des RKI werden auf [Zenodo.org](https://zenodo.org), [GitHub.com](https://github.com), [OpenCoDE](https://gitlab.opencode.de) und [Edoc.rki.de](https://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](#) .

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.