

# COVID-SARI-Hospitalisierungsinzidenz

---

[Robert Koch-Institute | RKI](#)

Nordufer 20  
13353 Berlin

[Kristin Tolksdorf](#), [Luise Goerlitz](#), [Walter Haas](#) und [Silke Buda](#)  
Fachgebiet 36 | Respiratorisch übertragbare Erkrankungen

---

## Zitieren

Tolksdorf K, Goerlitz L, Haas W und Buda S (2023): COVID-SARI-Hospitalisierungsinzidenz, Berlin:Zenodo. DOI: [10.5281/zenodo.7993899](https://doi.org/10.5281/zenodo.7993899)

## 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 sind essenziell für die Einschätzung der epidemiologischen Lage und die Anpassung der Maßnahmen während der COVID-19-Pandemie.

## Administrative und organisatorische Angaben

Im Datensatz "COVID-SARI-Hospitalisierungsinzidenz" wird die wöchentliche Inzidenz der Hospitalisierungen wegen einer neu aufgetretenen schweren Atemwegsinfektion (SARI) mit COVID-19 pro 100.000 Einwohnerinnen und Einwohner bereitgestellt.

Die zugrundeliegenden Daten werden vom Datenzentrum der HELIOS Kliniken GmbH an das Robert Koch-Institut (RKI) durch das übermittelt. Das Datenzentrum erhält die Daten von den teilnehmenden Krankenhäusern. Die Konzeptionierung der Datenerhebung, das Datenmanagement, die weitere Validierung der Daten und die fachliche Bewertung der Ergebnisse erfolgen im [Fachgebiet 36 | Respiratorisch übertragbare Erkrankungen](#) des RKI.

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

## Entstehungskontext

Die syndromische Surveillance schwerer akuter respiratorischer Infektionen (SARI) informiert über die aktuelle Krankheitsschwere und -häufigkeit. Dafür werden zeitnah erregerübergreifend akute Atemwegsinfektionen anhand von Symptomen bzw. den entsprechenden ärztlichen Diagnosen direkt an das RKI berichtet. Die syndromische Surveillance beruht auf systematisch und strukturiert erfassten Daten aus einer Stichprobe z.B. aus Krankenhäusern (Sentinel).

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.

## Datenerhebung

Derzeit werden Daten aus aktuell ca. 70 Sentinelkliniken mit einer bundesweiten Abdeckung von 5-6 % aller in Deutschland hospitalisierten Patient:innen erhoben. In den teilnehmenden Krankenhäusern werden fallbezogene Daten über aufgenommene Patient:innen routinemäßig erfasst. Die Daten aus diesen Krankenhäusern werden zentral im Datenzentrum im Rahmen des Qualitätsmanagements validiert. Von dort aus werden anonymisierte validierte fallbezogene Datensätze von neu im Krankenhaus aufgenommenen Patient:innen an das Robert Koch-Institut übermittelt ([Buda et al., 2017](#)).

Die übermittelten Daten teilen sich in zwei Datensätze auf: Der erste Datensatz (Datensatz 1) beschreibt die Gesamtzahl der Krankenhausaufenthalte innerhalb der meldenden Sentinel-Standorte und enthält Aufnahme- und Entlassungsdatum, Alter (in Jahren), Geschlecht und die ersten beiden Ziffern der 5-stelligen Postleitzahl eines jeden stationären Krankenhausaufenthalts. Dieser Datensatz enthält auch Informationen über die Dauer des Krankenhausaufenthalts, die Dauer des Aufenthalts auf der Intensivstation und die Art der Entlassung. Der zweite Datensatz (Datensatz 2) beschreibt die Untergruppe der Patient:innen, bei denen als Haupt- oder Nebendiagnose ein beliebiger ICD-10-Diagnosecode aus [Kapitel X \(Krankheiten des Atmungssystems J00 - J99\)](#) angegeben ist. Dieser zweite Datensatz enthält zusätzlich auch Informationen über die Einweisungsdiagnose, die primäre und alle sekundären Entlassungsdiagnosen, die Dauer der Beatmung, die ersten drei Ziffern der 5-stelligen Postleitzahl der Patient:in und die Hauptabteilung im Krankenhaus, in der die Patient:in behandelt wurde.

## 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 mit einer schweren akuten Atemwegserkrankung und COVID-19 im Krankenhaus behandelt wurden (COVID-SARI-Hospitalisierungsinzidenz).

Hierbei werden neu im Krankenhaus aufgenommene Patient:innen erfasst, die einen ICD-10-Code für SARI (ICD-10-Codes J09 bis J22) in der DRG-Haupt- oder Nebendiagnose sowie eine COVID-19-Diagnose (U07.1) erhalten haben, einschließlich noch hospitalisierter Personen. Die Berechnung der wöchentlichen COVID-SARI-Hospitalisierungsinzidenz erfolgt 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](https://doi.org/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](https://doi.org/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 robuste Aussagen zur Krankheitslast schwerer akuter Atemwegserkrankungen mit COVID-19.

## Aufbau und Inhalt des Datensatzes

---

Der Datensatz enthält die wöchentliche COVID-SARI-Hospitalisierungsinzidenz erhoben aus dem ICOSARI-Sentinel. 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:

- Berichtswoche des RKI
- wöchentliche Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen mit COVID-19

Die Daten werden dienstags im Rahmen der wöchentlichen Berichterstattung ausgewertet. Das bedeutet, dass alle bis dahin an das RKI übermittelten 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	Berichtswoche des RKI im ISO-8601 Format
agegroup	Text	60-79 , 80+ , 00+	Altersgruppen in Jahren, 00+ gibt die Gesamtinzidenz über alle Altersgruppen an.
icosari_covid19_incidence	Rationale Zahl		Hospitalisierungsinzidenz schwerer akuter respiratorischer Erkrankungen mit COVID-19

### Formatierung der Daten

Die Daten sind im Datensatz als kommaseparierte .csv Datei enthalten. Der verwendete Zeichensatz der .csv Datei ist UTF-8. Trennzeichen der einzelnen Werte ist ein Komma ",".

- Zeichensatz: UTF-8
- .csv Trennzeichen: Komma ","

## 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 Metadaten-Ordner 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](#)

## Hinweise zur Nachnutzung der Daten

---

Offene Forschungsdaten des RKI werden auf [github.com](#), [zenodo.org](#) und [edoc.rki.de](#) bereitgestellt:

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

## Lizenz

Der Datensatz "COVID-SARI-Hospitalisierungsinzidenz" ist lizenziert unter der [Creative Commons Namensnennung 4.0 International Public License](#) | [CC-BY 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/legalcode.de)

Die im Datensatz bereitgestellten Daten sind, unter Bedingung der Namensnennung des Robert Koch-Instituts als Quelle, frei verfügbar. Das bedeutet, jede:r 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.