Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland

Robert Koch-Institut | RKI Nordufer 20 13353 Berlin

Fachgebiet 25 | Körperliche Gesundheit

Hannelore Neuhauser (Projektleitung)
Christina Poethko-Müller
Giselle Sarganas Margolis
Roma Thamm

Fachgebiet 34 | HIV/AIDS und andere sexuell oder durch Blut übertragbare Infektionen Marlow Zimmermann

Fachgebiet 24 | Gesundheitsberichterstattung

Franziska Prütz

ZfKD | Zentrum für Krebsregisterdaten

Nina Buttmann-Schweiger Julia Fiebig

Neuhauser H, Buttmann-Schweiger N, Fiebig J, Poethko-Mülle C, Prütz F, Sarganas Margolis G, Thamm R und Zimmermann M (2022):: Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland, Berlin: Zenodo. DOI: 10.5281/zenodo.7043025.

Informationen zum Datensatz und zum Entstehungskontext

Die seit 2019 auftretende Infektionskrankheit COVID-19, hervorgerufen durch das neuartige SARS-CoV-2-Virus, führte zu gesundheitspolitischen und gesamtgesellschaftlichen Herausforderungen. Um geeignete Maßnahmen zur Eindämmung der Pandemie ergreifen zu können und neue Erkenntnisse über die Pandemie zu gewinnen, gibt es vermehrt Forschungsbedarfe zu COVID-19. Ein Ansatzpunkt hierfür sind die gewonnenen Blutproben von infizierten sowie von nicht infizierten Personen, die in Laboren auf Antikörper gegen das SARS-CoV-2-Virus getestet und analysiert werden. Sie geben Aufschluss über den Anteil der Bevölkerung, der bereits eine Infektion mit SARS-CoV-2 durchgemacht hat, und schließen dabei nicht erkannte Infektionen (Untererfassung) ein.

Das Projekt "Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland" (SERO-OBS Corona) gibt eine Übersicht zu Antikörper-Studien (sogenannte seroepidemiologische Studien) in Deutschland. Die seroepidemiologischen Studien basieren auf Blutproben von Bürgerinnen und Bürgern, die zu unterschiedlichen Zeitpunkten der Pandemie auf Antikörper gegen das SARS-CoV-2-Virus getestet wurden. Dabei sollen z. B. folgende Fragen beantwortet werden: Wie ist die Häufigkeit von SARS-CoV-2-Infektionen in verschiedenen Bevölkerungsgruppen? Wie hoch ist der Untererfassungsfaktor, der zeigt, wie viel Mal mehr Infektionen im Vergleich zu den bislang bekannten (gemeldeten) Fällen aufgetreten sind? In dem

vorliegenden Projekt werden in Deutschland durchgeführte seroepidemiologische Studien zu SARS-CoV-2 seit dem Frühjahr 2020 über systematische Recherchen in Studienregistern, Literaturdatenbanken einschließlich Vorveröffentlichungen sowie Medienberichten fortlaufend indentifiziert und Studieninformationen sowie Ergebnisübersichten verfügbar gemacht.

Die Ergebnisse des Projektes SERO-OBS-Corona werden auf der Webseite www.rki.de/covid-19-ak-studien, auf Deutsch, sowie der Webseite www.rki.de/covid-19-serostudies-germany, auf Englisch, bereitgestellt und regelmäßig aktualisiert.

Administrative und organisatorische Angaben

Im Projekt "Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland" (SERO-OBS Corona) wirken verschiedene Fachgebiete der Abteilung 2 | Epidemiologie und Gesundheitsmonitoring und Abteilung 3 | Infektionsepidemiologie des Robert Koch-Instituts mit. Wir bedanken uns bei dem Team von SeroTracker (serotracker.com) der Universität Calgary für hilfreiche Diskussionen und Beiträge zur Operationalisierung der Studiendatenbank sowie für die initiale Programmierung der interaktiven Graphiken.

Inhaltliche Fragen bezüglich der Datenextraktion können direkt an das Team des Projektes SERO-OBS-Corona unter COVID19_AK@rki.de gestellt werden. Die Veröffentlichung der Daten sowie das Qualitätsmanagement der (Meta-)Daten erfolgen durch das Fachgebiet MF 4 | Forschungsdaten- und Infromationsmanagement. Fragen zum Datenmanagement und zur Publikationsinfrastruktur können an das Open Data Team des Fachgebiets MF4 unter OpenData@rki.de gerichtet werden.

Das Projekt "Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland (SERO-OBS Corona)" wird durch das Bundesministerium für Gesundheit gefördert.

Inhalt und Aufbau des Forschungsprojekts

Seroepidemiologische Studien sind seit Beginn der COVID-19-Pandemie von hohem Interesse. Sie geben Aufschluss über den Anteil der Bevölkerung, der bereits eine Infektion mit SARS-CoV-2 durchgemacht hat, und schließen dabei nicht erkannte Infektionen (Untererfassung) ein.

Diese Studien werden seit dem Frühjahr 2020 über systematische Recherchen in Studienregistern, Literaturdatenbanken einschließlich Vorveröffentlichungen sowie Medienberichten fortlaufend identifiziert. Die übergeordnete Suchstrategie des RKI zum neuartigen Coronavirus ist in der Fachzeitschrift GMS der Arbeitsgemeinschaft für Medizinisches Bibliothekswesen veröffentlicht (Erling J, Heldt K, Sturm J, 2021).

Erling J, Heldt K, Sturm J. Lessons learned from the pandemic – ein Praxisbericht aus der Bibliothek des Robert Koch-Instituts. GMS Med Bibl Inf. 2021;21(1-2):Doc05. DOI: 10.3205/mbi000494

Es erfolgt eine kritische Bewertung der Studien, die Extraktion aggregierter Daten aus bereits vorliegenden Studien aus Deutschland sowie deren Aufbereitung für Analysen, für eine interaktive Karte, für ein Dashboard und für die Aktualisierung eines Factsheets zu den Ergebnissen seroepidemiologischer SARS-CoV-2-Studien in der Allgemeinbevölkerung in Deutschland. Weiterhin werden Recherchen und Analysen zur Untererfassung und Infection Fatality Rate durchgeführt. Die Ergebnisse des Projektes SERO-OBS-Corona werden auf der Webseite www.rki.de/covid-19-ak-studien auf Deutsch sowie der Webseite www.rki.de/covid-19-serostudies-germany auf Englisch bereitgestellt und regelmäßig aktualisiert.

Das Projekt "Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland" (SERO-OBS Corona) gibt somit einen Überblick über in Deutschland durchgeführte seroepidemiologische Studien zu SARS-CoV-2. Folgende Veröffentlichungen sind im Projekt entstanden:

Neuhauser H, et. al : Seroepidemiologische Studien zu SARS-CoV-2 in Stichproben der Allgemeinbevölkerung und bei Blutspenderinnen und Blutspendern in Deutschland – Ergebnisse bis August 2021 Epid Bull 2021;37:3 -12 | DOI: 10.25646/8999

Thamm, R., et al. SARS-CoV-2-Seroprävalenz bei Kindern und Jugendlichen in Deutschland – ein Überblick. Bundesgesundheitsblatt 64, 1483–1491 (2021). DOI: 10.1007/s00103-021-03448-0

Poethko-Müller C, et al.: Studien zur Seroprävalenz von SARS-CoV-2 in Deutschland und international. JoHM 2020;5(S(4)):2-16. DOI: 10.25646/7023

Neuhauser H, et al.: Ergebnisse seroepidemiologischer Studien zu SARS-CoV-2 in Stichproben der Allgemeinbevölkerung und bei Blutspenderinnen und Blutspendern in Deutschland (Stand 3.12.2020). Epid Bull 2020;50:3-6. DOI: 10.25646/7728

Robert Koch-Institut (2021) SARS-CoV-2-Seroprävalenz in der Allgemeinbevölkerung in Deutschland. Stand: 23. November 2021. DOI: 10.25646/9288

Robert Koch-Institut (2022) SARS-CoV-2-Seroprävalenz in der Allgemeinbevölkerung in Deutschland – Aktualisierung Juni 2022. https://edoc.rki.de/handle/176904/9849.2

Robert Koch-Institut (2022) SARS-CoV-2-Seroprävalenz in der Allgemeinbevölkerung in Deutschland – Aktualisierung September 2022. DOI: 10.25646/9693.3

Auswahl und Beschreibung des Untersuchungsgegenstandes

Informationen und Ergebnisse epidemiologischer Studien zur Seroprävalenz von SARS-CoV-2 in Zufallsstichproben der Allgemeinbevölkerung, bei Blutspenderinnen und Blutspendern in Deutschland und in besonderen Bevölkerungsgruppen wurden systematisch recherchiert und Eckdaten bereitgestellt.

Methoden, Instrumente und Verlauf der Datenerhebung

Recherche und Suchstrategie: Seroepidemiologische Studien zu SARS-CoV-2 in Deutschland, die als peerreviewte Publikation (mit wissenschaftlicher Begutachtung) in Literaturdatenbanken auffindbar sind (PubMed und Embase). Ebenso gesucht wird nach Manuskripten, die vor dem Peer-Review auf sogenannten Preprint-Servern veröffentlicht wurden (medRxiv, bioRxiv, arXiv, ChemRxiv, preprints.org, ResearchSquare, Social Science Research Network (SSRN)).

Die definierten Suchbegriffe im Titel der Publikationen aus Funden einer ersten Suche nach SARS-CoV-2 und verwandten Begriffen sind "sero OR antibod OR immune OR immunity OR immunology OR fatality rate OR population-based OR cohort study OR dried blood OR test strategy".

Auch Berichte und Meldungen der WHO, des European Centre for Disease Prevention and Control (ECDC) und des Centers for Disease Control and Prevention (CDC) sowie Studienregister (Deutsches Register Klinischer Studien, ClinicalTrials.gov) wurden durchsucht und die Medienberichterstattung verfolgt.

Inhalt und Aufbau des Datensatzes

Der Datensatz enthält Ergebnisse und Informationen aus seroepidemiologischen Studien zu SARS-CoV-2 in Deutschland. Zusätzlich werden grundlegende Metadaten sowie die Datensatzdokumentation bereitgestellt. Im Datensatz enthalten sind:

- Tabellen mit Studieninformationen und Ergebniszusammenfassungen
- Datensatzdokumentation in deutscher Sprache
- Lizenz-Datei mit der Nutzungslizenz des Datensatzes in Deutsch und Englisch
- Metadaten-Dateien mit Kontextinformationen

Daten des Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland

Die Daten des Projekts "Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland" (SERO-OBS Corona) sind im Hauptverzeichnis, in deutscher und englsicher Spache, in den Datenformaten .csv, .json und .xlsx abrufbar:

```
SERO-OBS_SARS-CoV-2_in_Deutschland.csv
SERO-OBS_SARS-CoV-2_in_Deutschland.json
SERO-OBS_SARS-CoV-2_in_Deutschland.xlsx
SERO-OBS_SARS-CoV-2_in_Germany.csv
SERO-OBS_SARS-CoV-2_in_Germany.json
SERO-OBS_SARS-CoV-2_in_Germany.xlsx
```

Die bereitgestellten Daten differenzieren nach übergeordneten Merkmalen:

- Informationen zur Studie und zum Studiendesign
- Ergebnis-Informationen
 - o zur Seroprävalenzangabe (Wert, KI, Adjustierung etc.),
 - zum Untererfassungsfaktor

Variablen und Variablenausprägungen

Eine Studie hat immer einen eindeutigen Namen, kann jedoch sowohl mehrere Publikationen umfassen, mehrere Studienorte haben etc. Bei Wechsel vom preprint zum Vollprint werden Information aktualisiert, wenn die spezifischen Ergebnisse (bspw. nach Alter, Gegend, Zeitraum) noch einmal neu berechnet wurden und voneinander abweichen. Die Einträge werden von einer Person angelegt, von einer zweiten Person qualitätsgesichert und bei Uneinigkeit in der Arbeitsgruppe besprochen.

Der Datensatz enthält folgende Variablen und mit ihren entsprechenden Ausprägungen:

Variable	Тур	Ausprägung	Erläuterung
title	Text	\$Name	Name aus Eigenangaben, DRKS, Studientitel

Variable	Тур	Ausprägung	Erläuterung
type	Text	Querschnitt, Längsschnitt/Prospektive Kohortenstudie, Quer- und selektiver Längsschnitt, Wiederholte Querschnitte	Querschnitt bei einmaliger Beprobung; Längsschnitt k wiederholter Beprobung de gleichen Teilnehmenden ü die Zeit; Quer- und selektiv Längsschnitt bei Auswahl Teilnehmenden zu Studienbeginn und selektiv Auswahl der weiteren zu beprobenden Teilnehmend aus dem Pool bereits Eingeschlossener; Wieder Querschnitte bei wiederho Beprobung an einem Ort/ü ein Rekrutierungsverfahrei mehreren Zeitpunkten.
start_date	Datum	JJJJ-MM-TT	Eindeutigstes Datum des Beginns der Probenentnah falls nur Monatsangabe verfügbar: wurde der 01. de Monats angegeben; bei Ar einer Zeitspanne wurde de des Monats gewählt
end_date	Datum	JJJJ-MM-TT	Eindeutigstes Datum des E der Probenentnahme, falls Monatsangabe verfügbar: der 01. des Monats angege bei Angabe einer Zeitspan wurde der 15. des Monats gewählt
organisation_\$	Text	\$Name	Name der Probennehmen Organisation
organisation_raw	Text	\$Name,\$Name,	Namen der Probennehmei Organisationen
location_\$	Text	\$Name	Kleinräumigste Angabe zu Beprobungsregion (z.B. St Kreis, Gemeinde) des Studienortes (Einzelwerte)
location_raw	Text	\$Name,\$Name,	Kleinräumigste Angabe zu Beprobungsregion (z.B. St Kreis, Gemeinde) des Studienortes

Variable	Тур	Ausprägung	Erläuterung
region	Text	Bundesweit, Baden- Württemberg, Bayern,, Thüringen, Multilokal	Name des Bundeslandes sein Wert für das gesamte Bundesgebiet und Bundes übergreifende Erhebunger (Mulitlokal)
registry_url	Text	https://www.NAME.de	Link zum Studienregistere (bspw. DRKS)
information_url	Text	https://www.NAME.de	Link zur Studienseite (bsp SARS-CoV-2: Antikörper- Studien des RKI)
resource_url_\$	Text	https://www.Name1.de	Angabe zum Link zu Publikation/Bericht, die im Rahmen der Studie veröffentlicht wurden. (Einzelwerte)
scope_regionality	Text	Bundesweit, Lokal/Regional	Der beschriebene räumlich Abdeckungsgrad des Schä (bspw. welche Region soll Schätzer repräsentieren.,Lokal/Regi innerhalb eines Bundeslar Bundesweit: mehrere Bundesländer
scope_population	Text	Allgemeinbevölkerung, Besondere Bevölkerungsgruppen, Allgemeinbevölkerung und besondere Bevölkerungsgruppen	Die Stichprobe repräsentie Allgemeinbevölkerung ode einen Eindruck der Seroprävalenz in Bevölkerungsgruppen, die Auswahl der Allgemeinbevölkerung darstellen. Bspw. Zielpopu einer Studie von Schülerin und Schülern im schulpflichtigen Alter eines Bundeslandes sollen die Seroprävalenz der Allgemeinbevölkerung die Alters wiedergeben; Lehrpersonal hingegen ste Zielpopulation eine besone Bevölkerungsgruppe dar.

Variable	Тур	Ausprägung	Erläuterung
scope_age	Text	Erwachsene, Kinder/Jugendliche, Mehrere	Die Auswahl erfolgt exklus Kinder und Jugendlich sin <=18 Jahre
subject_frame	Text		Auswahl des Rahmens, in die Stichprobe gezogen w für die das Ergebnis zur Seroprävalenz abgebildet Die Auswahl gibt die Grup wieder, die in der Studie al spezifischsten beschriebel wurde (bspw. Bewohnende Altenheimen anstatt Wohnbevölkerung)
subject_information	Text		Verknüpfung von Informationaus dem Stichprobenrahm und dem Studienort und geweiteren relevanten Charakteristika der Stichprim Sinne einer Kurzbeschreibung.
subject_vaccination	Text	Seroprävalenz in allen Teilnehmenden, Anteil der Seropositiven in Ungeimpften / in Teilnehmenden vor 2021, Anteil der Seropositiven in Geimpften, Anteil der Seropositiven in Ungeimpften	Spezifikation des Impfstatu Teilnehmenden zu den in d Grafiken dargestellten Ergebnissen.

Variable	Тур	Ausprägung	Erläuterung
sample_design	Text	Nicht-zufällige Ziehung, Blutspende, Vollerhebung der Grundgesamtheit, Kohortenstichprobe, Nicht-Einwohnermeldeamts- Zufallsstichprobe, Einwohnermeldeamts- Zufallsstichprobe, Mehrstufige, teilweise zufällige Ziehung	Nicht-zufällige Ziehung/convenience sam Nicht zufällige Teilnehmendenrekrutierun einfach verfügbaren Grupp Blutspende: Anonymisierte Restproben aus Blutspend werden für die Studie verwendet. Vollerhebung: Person innerhalb einer Population wird in die Stude eingeschlossen. Kohortenstichprobe: Sticht aus einer bestehenden Kohortenstudie in der Allgemeinbevölkerung. Nic Einwohnermeldeamts- Zufallsstichprobe: Nicht au Basis einer Einwohnermeldeamtsstich zufällig ausgewählte Teilnehmende. Einwohnermeldeamts- Zufallsstichprobe: Auf Bas einer Einwohnermeldeamtsstich zufällig ausgewählte Teilnehmende. Mehrstufige teilweise zufällige Ziehung Mehrstufige Stichprobenziehung, nicht allen Stufen zufällig.
sample_analysis_unit	Text	Teilnehmende, Haushalte mit Teilnehmenden;, Samples von Teilnehmenden, Arztpraxen mit Teilnehmenden	Einheit der Stichprobenzie
age_max	Ganze Zahl		Höchstes Alter von in die § eingeschlossenen Teilnehmenden

Variable	Тур	Ausprägung	Erläuterung
age_min	Ganze Zahl		Geringstes Alter von in die Studie eingeschlossenen Teilnehmenden
sample_size_obtained	Ganze Zahl		Anzahl der in die Studie eingeschlossenen Teilnehmenden
sample_size_rounded	Ganze Zahl		Gerundet Wert von sample_size_obtained
analysis_data_source	Text	Vollblut, Kapillarblut, Blutplasma, Kein Blut, Sputum	Methode der Blutabnahme Antikörperbestimmung
analysis_target	Text	anti -S1 domain , anti - RBD domain , anti Nucleocapsid , multiple	Zielstruktur des Virus, geg sich die in der Studie bestimmten Antikörper rich
analysis_isotype	Text	Gesamt Ig-Antikörper, IgG, Mehrere	Angaben zur Antikörperkla auf die sich die berichtete Seroprävalenz bezieht.
analysis_test	Text	ELISA, Elektrochemolumineszenz (ECL)-Immunassay, LIPS, andere	Angaben zum laboranalyti Testverfahren, mit dem die Antikörperbestimmung durchgeführt wurde .
analysis_test_manufacturer_\$	Text	\$Name	Hersteller des Testassays. (Einzelwerte)
analysis_test_manufacturer_raw	Text	\$Name, Name, Name,	Hersteller des Testassays.
results_publication_url	Text	https://www.NAME.de	Referenz auf Quelle, in de Information zum jeweiliger Ergebnis (bspw. Seropräva Dunkelziffer) vorgehalten v
results_publication_date	Datum	JJJJ-MM-DD	Veröffentlichungsdatum de Referenz auf Quelle, in de Information zum jeweiliger Ergebnis (bspw. Seropräva Dunkelziffer) vorgehalten v

Variable	Тур	Ausprägung	Erläuterung
results_publication_type	Text		Typ der Quelle, in der die Information zum jeweiliger Ergebnis (bspw. Seropräva Dunkelziffer) vorgehalten v (bspw. Pre-Print, Vollveröffentlichung als Pereviewte Publikation).
results_seroprevalence_ adjustment	Text	Nicht adjustiert, Bevölkerungsgewichtet, Adjustiert für Test- Güte, Adjustiert für Test-Güte und bevölkerungsgewichtet	Adjustierung der Seropräv für Faktoren wie Test-Güte Bevölkerung: Bei mehrere Varianten wurde die höchs Adjustierung angegeben.
results_seroprevalence_ sensitivity	decimal		Testgüteeigenschaften des Antikörpertests, die für die Analysen herangezogen wurden.
results_seroprevalence_ specificity	decimal		Testgüteeigenschaften des Antikörpertests, die für die Analysen herangezogen wurden.
results_seroprevalence_ positive	decimal		Seroprävalenzangaben mi jeweils höchsten verfügba bzw. aus den Studienanga errechenbare Adjustierungsstufe.
results_seroprevalence_ positive_95pct_lower	decimal		Unteres 95% Konfidenzint zur Seroprävalenz
results_seroprevalence_ positive_95pct_upper	decimal		Oberes 95% Konfidenzinte zur Seroprävalenz
results_seroprevalence_ information	Text		Verknüpfung von Informatic zur Seroprävalenz, die bsp den Studienort und den Probezeitraum beinhalten ggfs. weiteren relevanten Charakteristika, die innerh einer Studie die ermittelter Seroprävalenzen beschrei

Variable	Тур	Ausprägung	Erläuterung
results_underreporting	Text		Faktor, der zeigt, wie viel N mehr Infektionen im Vergle zu den bis zum Studienmittelpunkt bekann (gemeldeten) kumulativen Fällen aufgetreten sind.
record_created	Datum	JJJJ-MM-DD	Zeitpunkt, zu dem ein neue Ergebnis aufgenommen w
record_modified	Datum	JJJJ-MM-DD	Zeitpunkt, zu dem ein bere bestehendes Ergebnis angepasst wurde (z.B. pee reviewte Vollpublikation ar eines Preprints)

Formatierung der Daten

Die Daten der Studie 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

Konexinformationen

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 "Observatorium serologischer Studien zu SARS-CoV-2 in Deutschland" 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.