COVID-19-Impfungen in Deutschland

Robert Koch-Institut | RKI Nordufer 20 13353 Berlin

FG 33 | Impfprävention

Anette Siedler (Fachliche Leitung)
Thorsten Rieck (Datenanalyse)
Annika Steffen (Datenanalyse)
Nita Perumal (Datenanalyse)
Nathalie Pomijalski (Dokumentation)

IT 4 | Development

Lei Mao (Projektleitung)
Hermann Claus (Technische Leitung)
Constantin Fischer (Datenmanagement)
Sven Schröder (Datenmanagement)
Daniel Wesseler (Datenmanagement)

MF 4 | Informations- und Forschungsdatenmanagement

Hannes Wuensche (Datenkuration)

Robert Koch-Institut (2022): COVID-19-Impfungen in Deutschland, Berlin: Zenodo. DOI:10.5281/zenodo.7073917.

Informationen zum Datensatz und Entstehungskontext

Die COVID-19-Impfung wurde als Wendepunkt in der Kontrolle der COVID-19-Pandemie angesehen und erfährt daher hohes Maß an öffentlicher Aufmerksamkeit. Alle Leistungserbringer von Impfungen sowie die Erfassung der Impfdaten sind in der Coronavirus-Impfverordnung festgelegt. Gemäß Impfverordnung sind alle Impfdaten täglich dem Robert Koch Institut (RKI) zu melden. Die Impfkampagne wurde in ihrem Verlauf um immer weitere Leistungserbringer erweitert, für die mehrere Meldeportale eingerichtet wurden, die in diesem Dokument beschrieben werden.

Ziel des Projekts "Digitales Impfquoten-Monitoring" (DIM) ist es, tagesaktuell und bundesweit die Impfdaten aus allen Meldeportalen zu erfassen und folgend aufbereitet darzustellen, um zeitnah Impfquoten und damit den Verlauf der COVID-19-Impfkampagne zu analysieren. Die Impfdaten werden am RKI weiterhin für die Analyse von Impfdurchbrüchen und der Wirksamkeit der Impfstoffe genutzt. Sie sind damit eine wichtige Grundlage für die Verantwortlichen bei Bund, Ländern und Kommunen für die Steuerung der Impfkampagne.

Projektauftraggeber ist das Bundesministerium für Gesundheit (BMG). Das Robert Koch-Institut übernimmt die Projektleitung. Projektzeitraum ist vom 01.11.2020 bis 30.11.2022.

Ein wesentlicher Bestandteil des DIM-Projekts ist die DIM-Anwendung. Diese Web-basierte Softwarelösung wurde zur Erfassung der Impfdaten insbesondere für die Leistungserbringer eingerichtet, die extra und zusätzlich zur medizinischen Regelversorgung an der COVID-Impfkampagne beteiligt sind.

An der Entwicklung und dem Betrieb der DIM-Web-Anwendung sind weitere Stakeholder beteiligt. Die Bundesdruckerei (Bdr) ist zuständig für die technische Entwicklung und den Betrieb der DIM-Web-Anwendung. Accenture GmbH (ACN) unterstützt das RKI bei der Konzeption, der Planung, Dokumentation und der Rolloutkoordination.

Die DIM-Anwendung besteht aus Eingabemasken/Dialogen (Frontend) und dahinter liegendem Datenbanksystem (Backend). Autorisiertes Personal der stationären Impfzentren, mobilen Impfteams und weiterer in der CoronalmpfVO ausgewiesener Leistungserbringer können über die Web-Anwendung die Impfdatensätze zur COVID-19-Impfung eingeben und über eine gesicherte Internetverbindung pseudonymisiert an die Bundesdruckerei übermitteln. Im Auftrag des RKI werden die Daten von der Bundesdruckerei zwischengespeichert und vom RKI mehrmals täglich abgerufen. Alternativ zur direkten Eingabe können Daten in einem vorgegebenen csv-Datenformat durch die DIM-Web-Anwendung übermittelt werden (csv-Upload Schnittstelle). Als dritte Möglichkeit können Datensätze über eine REST-API Schnittstelle aus einem Fremdsystem in das Datenbanksystem übertragen werden.

Die DIM-Datenbank am RKI speist sich aus weiteren Datenportalen, die jedoch technisch nicht vom RKI umgesetzt wurden (s. Beschreibung im Abschnitt Zusammensetzung der Datenquellen).

Die nachfolgende Darstellung zeigt schematisch den Datenfluss von Erfassung der Daten bei den impfenden Stellen bis hin zur Publikation und Weitergabe der aggregierten Daten durch das RKI.



Rollenbesetzung im RKI

Die fachlich-technische Projektleitung liegt im Fachgebiet IT 4 | Development des RKI. Hier erfolgt die Verarbeitung und Aufbereitung der vorliegenden Rohdaten.

Die fachlich-epidemiologische Projektleitung liegt im Fachgebiet 33 | Impfprävention des RKI. Hier erfolgt die epidemiologische Datenanalyse.

Inhaltliche Fragen zum Digitalen Impfquoten-Monitoring und dem Impf-Fortschritt können direkt an info@rki.de gerichtet werden.

Die Veröffentlichung der Daten, die Datenkuration sowie das Qualitätsmanagement der (Meta-)Daten erfolgen durch das Fachgebiet MF 4 | Forschungsdatenmanagement. Fragen zum Datenmanagement und zur Publikationsinfrastruktur können an das Open Data Team des Fachgebiets MF4 unter OpenData@rki.de gerichtet werden.

Zusammensetzung der Datenquellen

Das DIM-Projekt enthält Daten über den Verlauf der COVID-19-Impfungen in Deutschland. Die hier veröffentlichten Impfdaten sind aggregierte Daten aus vier Datenquellen. Neben dem Meldeweg über die DIM-Anwendung werden zur Erfassung der Impfquote auch Impfdaten aus Arztpraxen über die Kassenärztliche Bundesvereinigung (KBV), die Kassenärztlichen Vereinigungen (KV) und die Privatärztlichen Verrechnungsstellen (PVS) erfasst. Insgesamt werden im Datensatz "COVID-19-Impfungen in Deutschland" somit verschiedenste Datenquellen zusammengeführt und ständig erweitert:

 Mit der DIM-Webanwendung werden die Impfdaten von den Leistungserbringern übermittelt, die außerhalb des Routineimpfsystems in die COVID19-Impfkampagne einbezogen sind.

Aus Impfzentren, mobilen Impfteams und einigen Krankenhäusern der Bundesländer werden seit dem 27.12.2020, von Betriebsärzt:innen und betriebsmedizinischen Diensten seit dem 07.06.2021, aus Gesundheitsämtern und weiteren Krankenhäusern seit Oktober 2021 und aus Apotheken seit dem 08.02.2022 jeweils täglich pseudonymisierte Daten mit der Web-Anwendung "Digitales Impfquotenmonitoring" übermittelt. Seit dem 25.01.2022 können alle Impfungen entsprechend ihrer tatsächlichen Dosis in der Impfserie von 1 bis 5 (auch rückwirkend) übermittelt werden. Bis zu diesem Zeitpunkt gab es impfstoffspezifische Besonderheiten, auf die weiter unten noch eingegangen wird.

• Über ein Meldeportal der Kassenärztliche Bundesvereinigung (KBV) übermitteln die impfenden Vertragsärzt:innen täglich einen aggregierten Kerndatensatz.

Von den Vertragsärzt:innen werden täglich aggregierte Daten an die KBV übermittelt und dort in zwei separaten Datenpaketen dem RKI bereitgestellt. Die Pakete beinhalten 1) die Anzahl durchgeführter Impfungen je Praxis und Impftag aufgeschlüsselt nach Impfstoff und Impfstoffdosis (verfügbar seit Impftag 10.03.2021) sowie 2.) die Anzahl der Impfungen aufgeschlüsselt nach Altersgruppe <18, 18-59 und 60+ Jahre und Impfstoffdosis (verfügbar seit Impftag 26.03.2021). Die beiden Datenpakete sind nicht miteinander verknüpfbar. Eine Zuordnung von Impfstoff und Altersgruppe ist mit diesen Daten nicht möglich. Seit Impftag 14.12.2021 wird die Altersgruppe <18 Jahre aufgeschlüsselt nach Altersgruppe 5-11 sowie 12-17 Jahre übermittelt. Seit dem 27.01.2022 sollten im KBV-Portal alle Impfungen entsprechend ihrer tatsächlichen Dosis in der Impfserie übermittelt werden können; seit dem 10.02.22 liegen Daten mit der Impfserie 4 vor. Besonderheiten gab es bei der Übermittlung von Impfungen mit Janssen-Impfstoff (s. weiter unten).

Eine regionale Zuordnung der Impfdaten ist ausschließlich nach der PLZ der Arztpraxis möglich.

 Quartalsweise Abrechnungsdaten der Kassenärztlichen Vereinigungen (KV) ersetzen rückwirkend die KBV-Daten.

Die Daten der Vertragsärzte aus dem KBV-Schnellmeldeportal werden sukzessive durch die quartalsweisen Abrechnungsdaten der Kassenärztlichen Vereinigungen (KV) ersetzt. In den KV-Daten ist

die regionale Zuordnung nach der Wohnort-PLZ der Geimpften sowie die Zuordnung von Impfstoff und Altersgruppe möglich.

Ab dem 22.06.2022 gilt: Die KBV-Schnellmeldungen mit Impfungen bis zum Ende des vierten Quartals 2021 (31.12.2021) wurden durch Daten der 17 KVen ersetzt, wobei folgende Einschränkungen gelten:

- Für den Zeitraum 14.12.-31.12.21 liegen die KV-Abrechnungsdaten der <18-Jährigen nur für die Gesamtaltersgruppe vor, nicht aber weiter differenziert nach 5-11 Jahre und 12-17 Jahre. Daher wurde eine Ersetzung der KBV-Daten durch die KV-Daten für diesen Zeitraum zwar für die Gesamtgruppe <18 Jahre vorgenommen, nicht jedoch für die beiden Einzelaltersgruppen 5-11 und 12-17 Jahre, für die weiterhin die KBV-Daten verwendet wurden. Somit können rund 35.000 Impfungen im Alter <18 Jahren nicht korrekt den beiden Einzelaltersgruppen zugewiesen werden.
- Für 128.289 übermittelte Impfungen der KV Brandenburg liegt weder eine Differenzierung nach Impfstoff noch nach Impfserie vor. Diese Impfungen blieben unberücksichtigt.
- Privatärzt:innen übermitteln täglich einen aggregierten Kerndatensatz über die Privatärztlichen Verrechnungsstellen (PVS).

Von den Privatärzt:innen werden täglich aggregierte Daten an die PVS übermittelt und dort in einem Datenpaket dem RKI bereitgestellt. Dieses Paket enthält die Anzahl durchgeführter Impfungen je Praxis und Impftag aufgeschlüsselt nach Impfstoff, Altersgruppe <18, 18-59 und 60+ Jahre und Impfstoffdosis (Daten verfügbar seit Impftag 07.06.2021). Seit Impftag 10.12.2021 wird die Altersgruppe <18 Jahre aufgeschlüsselt nach Altersgruppe 5-11 sowie 12-17 Jahre übermittelt. Seit Impftag 19.01.2022 können alle Impfungen entsprechend ihrer tatsächlichen Dosis in der Impfserie übermittelt werden, seit dem 10.02.2022 liegen Daten zur Impfserie 4 vor. Auch in diesen Daten ist eine regionale Zuordnung der Impfdaten ausschließlich nach der PLZ der Arztpraxis möglich. Die Daten der Privatärzt:innen liegen nicht noch einmal in anderen Datenbeständen wie zum Beispiel in Abrechnungsdaten feiner aufgeschlüsselt vor und können daher nicht ersetzt werden.

Impfmeldungen mit Janssen-Impfstoff

Besonderheiten gab es bei der Meldung von Impfungen mit dem Janssen-Impfstoff (verfügbar seit Mai 2021). Für diesen Impfstoff war zunächst eine Impfstoffdosis als ausreichend für die Grundimmunisierung angesehen worden. Seit Oktober 2021 sind von der Ständigen Impfkommission Optimierungsimpfungen nach einmaliger Janssen-Impfung mit mRNA-Impfstoffen empfohlen. Mit Wirkung vom 15.01.2022 sind beim Janssen-Impfstoff generell zwei Impfungen für die Grundimmunisierung vorgesehen. Alle Meldesysteme (DIM, PVS, KBV) wurden im Laufe des Januar 2022 dahingehend angepasst, dass alle Impfungen (inklusive Janssen-Impfungen) entsprechend der tatsächlichen Folge in der Impfserie gemeldet und gewertet werden können. Die Meldung von Janssen-Impfungen und die Umsetzung der Änderungen erfolgte in den einzelnen Portalen unterschiedlich. Seit dem 29.04.2022 werden alle Daten entsprechend ihrer tatsächlich gemeldeten Impfserie publiziert. Auch dabei gibt es jedoch Unterschiede.

Im **DIM-Portal** wurden Janssen-Impfungen bis zum 24.01.2022 ausschließlich mit der Impfserie 1 gemeldet und in der Analyse zusätzlich der vollständigen Grundimmunisierung (Impfserie 2) hinzugerechnet. Optimierungsimpfungen waren dementsprechend bis 24.01.22 als Auffrischimpfung (Impfserie 3) zu dokumentieren. Ab 25.01.22 konnten hier alle Impfungen entsprechend ihrer tatsächlichen Reihenfolge dokumentiert und gemeldet werden. Die Datenanalyse und -Publikation wurde bez. der Janssen-Impfung zum 29.04.2022 wie folgt umgestellt: Einmalige Janssen-Impfungen zählen nur noch als Erstimpfung (Impfserie 1; das Hinzurechnen zur Impfserie 2 entfällt), Optimierungsimpfungen wurden soweit möglich in Zweitimpfungen umbewertet.

In den aggregierten Daten der Ärzteschaft ist diese Umbewertung jedoch nicht möglich;
Optimierungsimpfungen können nicht als solche identifiziert werden. Im KBV-Datenpaket mit
Impfstoffangabe wurden bis zum Impftag 05.01.2022 alle Impfungen mit Janssen-Impfstoff mit der Impfserie 2
übermittelt. Ab dem Impftag 06.01.2022 wurden Impfungen mit dem Janssen-Impfstoff auch mit der Impfserie
3 übermittelt, seit dem 27.01.2022 ist für Janssen-Impfungen die Angabe der Impfserie 1 möglich. Alle bis
zum 26.01.2022 im KBV-Portal übermittelten Janssen-Impfungen mit der Impfserie 2 werden der Impfserie 1
zugeordnet, da diese Impfungen mit hoher Wahrscheinlichkeit tatsächliche Erstimpfungen waren. Für die
Impfquotenbestimmung werden diese Impfungen sowohl bei den mindestens einmal Geimpften als auch –
der Übermittlung entsprechend - bei den Grundimmunisierten berücksichtigt. Alle ab dem 27.01.2022
übermittelten Impfungen mit Janssen-Impfstoff werden mit ihrer tatsächlich übermittelten Impfserie
ausgewiesen.

Die oben genannte Ersetzung der KBV-Daten durch **Abrechnungsdaten der KVen** ändert am prinzpellen Vorgehen nichts: Aus der impfstoffspezifischen Abrechnungsziffer für den Janssen-Impfstoff lässt sich keine Impfserie ableiten. Lediglich die Zuordnung nach Impfstoff und Altersgruppe kann durch die KV-Daten sukzessive vorgenommen werden.

Im **PVS-System** wurden bis zum Impftag 18.01.22 alle Impfungen mit dem Janssen-Impfstoff als Impfserie 1 übermittelt. Diese Impfungen wurden bis zur Umstellung der Datenanalyse und -Publikation am 29.04.2022 zusätzlich auch zur Impfserie 2 (Abschluss der Grundimmunisierung) hinzugerechnet; diese Hinzurechnung entfällt ab dem 29.04.22 auch rückwirkend. Seit dem 19.01.22 ist hier eine Übermittlung aller Impfungen entsprechend der tatsächlichen Impfserie möglich.

Aufbereitung der übermittelten Daten

Die Rohdaten der einzelnen Datenquellen werden für den Auswertungs-Datensatz aufbereitet. Für alle Datenquellen werden die folgenden Schritte durchgeführt:

- 1. Grundlegende Validitätsprüfung der übermittelten Datensätze
- 2. Einheitliche Benennung der Impfstoffe
- 3. Zuweisung der BundeslandID oder LandkreisID des Impfortes
- 4. Ausschluss von Datensätzen ohne Angabe des Impfzentrums sowie bei einer Altersangabe kleiner "2" (nur DIM)
- 5. Ausschluss von Impfungen mit Novavax vor dem 25.02.2022 (da Novavax in Deutschland erst seit diesem Tag zur Verfügung steht)
- 6. Erkennung von Sammelpatienten:innen und Auswahl der aktuellsten Datensätze je Pseudonym/Sammelpatienten:innen und Impfdatum (nur DIM)
- 7. Prüfung auf inhaltliche Konsistenz und ggf. Korrekturen (nur DIM)
- 8. Filterung auf Impfdaten im Zeitraum vom 27.12.2020 bis zum vorhergehenden Tag

Die folgende Beschreibung des Vorgehens bezieht sich allein auf die Daten aus dem DIM-Portal:

Korrektur von mehrfachen Meldungen

Den übermittelten DIM-Daten wird ein personenbezogenes Pseudonym zugeordnet. Beispielsweise, durch die spätere Korrektur von Meldungen kommt es dazu, dass zu einem personenbezogenen Pseudonym mehrere, sich widersprechende Einträge vorliegen. Das ist z.B. der Fall, wenn fälschlicherweise eine

Erstimpfung als Zweitimpfung gemeldet wird und später zur Erstimpfung korrigiert wird. Im Datensatz wird dann zunächst die Zweitimpfung und, nach gemeldeter Korrektur, die Erstimpfung ausgewiesen werden.

Erkennung von Sammelpatient:innen

Ein weiterer Fall von mehrfach vorkommenden Pseudonymen sind sogenannte Sammelpatient:innen, d.h. unterschiedliche Personen, denen dasselbe Pseudonym zugeordnet wird. Liegen Datensätze mit gleichem Pseudonym aber Unterschieden in der angegebenen Impfstelle und der Postleitzahl der Person vor und kommt keines dieser Merkmalsausprägungen in einer anderen Kombination dieser Merkmale vor, ist davon auszugehen, dass es sich um Sammelpatient:innen, d.h. verschiedene Personen, handelt. Die Annahme wird auch dann getroffen, wenn nach der Korrektur vor mehrfachen Meldungen Unterschiede in der Postleitzahl der Person oder der angegebenen Impfstelle zu unterschiedlichen Tagen vorliegen. Im folgenden Aufbereitungsprozess werden alle zu einem erkannten Sammelpatient:innen gehörigen Datensätze wie im Falle eines neuen Pseudonyms behandelt.

Regeln zur Herstellung inhaltlicher Konsistenz

Zur Herstellung inhaltlicher Konsistenz wird ein Mindestabstand zwischen Impfungen, Datumsgrenzen für einzelne Impfserien und Regeln für Einträge mit gleichem Personen-Pseudonym/erkannten Sammelpatient:innen verwendet. Die Einträge mit gleichem Personen-Pseudonym/erkannten Sammelpatient:innen heißen konsistent, wenn aufsteigende Impfserien mit aufsteigendem Impfdatum vorliegen, der Mindestabstand und Datumsgrenzen eingehalten sind, auf eine Erstimpfung mit Janssen keine erste Auffrischimpfung mit Impfdatum vor dem 26.01.2022 direkt folgt sowie keine unbekannte Impfserie vorliegt. Für den Mindestabstand zwischen zwei Impfungen beliebiger Impfserie werden 14 Tage angenommen. Impfungen mit Serie 2 werden ab dem 15.01.2021 und erste bzw. zweite Auffrischimpfungen ab dem 01.06.2021 bzw. 01.09.2021 als konsistent angesehen. Der Hinweis "aktuellste" bezieht sich dabei auf das Datum der Datenerfassung, nicht das Impfdatum.

Konsistenzregeln für Einträge mit gleichem Personen-Pseudonym/erkannten Sammelpatient:innen

- Wenn die Einträge konsistent sind, so werden diese ohne Korrekturen ausgewiesen.
- Wenn es mindestens zwei Einträge gibt und alle Einträge innerhalb von 13 Tagen liegen, dann wird nur der aktuellste Eintrag behalten, der die Datumsgrenzen wahrt und eine bekannte Impfserie hat.
- Sonst verwerfe alle Auffrischimpfungen, die vor ihrer Datumsgrenze liegen, und wende für jeden Eintrag in aufsteigender Reihenfolge gemäß des Impfdatums die folgenden Schritte an, wobei für eine unbekannte Impfserie beim aktuellen Eintrag die um eins erhöhte Impfserie der letzten ausgewiesenen Impfung bzw. eine 1 angenommen wird:

Schritt	Kriterium	Aktion
1	Der Abstand des aktuellen Eintrags zur letzten ausgewiesenen Impfung ist kleiner als der Mindestabstand.	Ja : Verwerfe den Eintrag Nein: weiter mit Schritt 2

Schritt	Kriterium	Aktion
2	Der aktuelle Eintrag ist eine erste Auffrischimpfung vor dem 26.01.2022 und die letzte ausgewiesene Impfung ist eine Erstimpfung mit Impfstoff Janssen.	Ja : Ändere die Impfserie zu 2 und weise die Impfung aus Nein: weiter mit Schritt 3
3	Der nächste Eintrag hat keine größere Impfserie, der Mindestabstand zu diesem ist eingehalten, die Impfserie der letzten ausgewiesenen Impfung ist mindestens 2 kleiner als die des nächsten Eintrags und das Impfdatum hält die Datumsgrenze für die um eins herabgesetzte Impfserie des Folgeeintrags ein.	Ja : Ändere die Impfserie zu der um eins herabgesetzten Impfserie des Folgeeintrags und weise die Impfung aus Nein: weiter mit Schritt 4
4	Der aktuelle Eintrag hat eine größere Impfserie als die letzte ausgewiesene Impfung und die Datumsgrenze ist eingehalten.	Ja : Weise die Impfung aus. Nein: weiter mit Schritt 5
5	Der aktuelle Eintrag hat keine größere Impfserie als die letzte ausgewiesene Impfung, die um eins höhere Impfserie der aktuellen Eintrags ist nicht größer als die maximal ausgewiesene Impfserie (4) und für diese Impfserie ist die Datumsgrenze eingehalten.	Ja : Ändere die Impfserie zu der um eins erhöhten Impfserie des aktuellen Eintrags und weise die Impfung aus Nein: verwerfe den aktuellen Eintrag

Wichtig ist, dass innerhalb des DIM Systems zwischen Impfdatum und dem Datum der Datenerfassung unterschieden wird. Meldungen als auch deren Korrekturen können sich auf das gleiche Impfdatum beziehen, das Datum der Datenerfassung unterscheidet sich ggf. jedoch, da die Korrektur später gemeldet wird.

Aufbau und Inhalt des publizierten Datensatzes

Der Datensatz enthält Daten über den Verlauf der COVID-19-Impfungen in Deutschland und die in der Datenverarbeitung unterstützenden Kontextmaterialien. Im Datensatz enthalten sind:

- Impfdaten mit tagesaktuellen Meldungen von COVID-19-Impfungen
- Datenschema der Impfdaten
- Archiv mit der Sammlung aller bisherigen Impfzahlentabellen
- Lizenz Datei mit der Nutzungslizenz des Datensatzes
- Datensatzdokumentation in deutscher Sprache
- Metadaten Datei zum Import in Zenodo

Formatierung der COVID-19 Impfdaten

Die Impfdaten sind im Datensatz als kommaseparierte .csv Dateien enthalten. Der verwendete Zeichensatz der .csv Dateien ist UTF-8. Datumsangaben sind im ISO8601 Standard formatiert.

· Zeichensatz: UTF-8

Datumsformat: ISO8601

· .csv Trennzeichen: Komma ","

Metadaten

Die bereitgestellten Daten sind mit Metadaten beschrieben und wissenschaftlich zitierbar, u.a. durch die Vergabe einer DOI durch Zenodo.org. Die für den Import in Zenodo bereitgestellten Metadaten sind in folgender Datei hinterlegt:

.zenodo.json

Die Dokumentation der einzelnen Metadatenvariablen ist unter https://developers.zenodo.org/#representation nachlesbar.

Neben der Beschreibung bibliographischer Metadaten, in der .zenodo.json, wird das Datenschema der Impfdaten bereitgestellt:

Kontextmaterialien/Datenschema_Deutschland_Bundeslaender_COVID-19-Impfungen.json Kontextmaterialien/Datenschema Deutschland Landkreise COVID-19-Impfungen.json

Im Datenschema aufgeführt sind die in den Impfdaten enthaltenen Variablen sowie deren Ausprägungen.

COVID-19 Impfdaten auf Ebene der Bundesländer

Die aktuellen COVID-19-Impfzahlen auf Ebene der Bundesländer sind im Hauptverzeichnis unter "Aktuell_Deutschland_Bundeslaender_COVID-19_Impfungen.csv" abrufbar und werden täglich überschrieben. Im Archivordner sind die täglichen Impfdaten der Bundesländer unter den Dateinamen "JJJJ-MM-TT_Deutschland_Bundeslaender_COVID-19-Impfungen.csv" abgelegt.

Aktuell_Deutschland_Bundeslaender_COVID-19-Impfungen.csv Archiv/JJJJ-MM-TT_Deutschland_Bundeslaender_COVID-19-Impfungen.csv

Die Impfdaten bilden einen tagesaktuellen Stand (08:00 Uhr) aller an das RKI gemeldeten Impfungen in Deutschland ab. Im Dateinamen repräsentiert die Sequenz "JJJJ-MM-TT" das Erstellungsdatum der Datei und gleichzeitig das Datum des enthaltenen Datenstands. "JJJJ" steht dabei für das Jahr, "MM" für den Monat und "TT" für den Tag der Erstellung bzw. des enthaltenen Datenstands. Die "Aktuell_Deutschland_Bundeslaender_COVID-19-Impfungen.csv" ist identisch mit dem neusten Datenstand des Archivs.

Variablen

Die Impfdaten differenzieren nach verschiedenen Merkmalen einer "Impfgruppe". Unter "Impfgruppe" wird die Zusammenfassung der Merkmalsausprägungen pro Impftag verstanden. Pro Eintrag bzw. Zeile ist eine eineindeutige Impfgruppe für ein Impfdatum abgebildet. Eine Impfgruppe umfasst in der Regel keine Einzelfälle. Jedoch ist es möglich, dass in einer Impfgruppe nur ein Fall enthalten ist. Eine Impfgruppe wird

grundlegend durch folgende Merkmale charakterisiert (in den Klammern finden sich die Variablen dieser Merkmale):

- Ort der Impfung (BundeslandId_Impfort)
- Impfung (Impfstoff, Impfserie)

Zusätzlich werden folgende Merkmale bzw. Variablen angegeben:

- Datum der Impfung (Impfdatum)
- Anzahl der Impfungen in der Gruppe (Anzahl)

Für jede Impfgruppe wird die tägliche Anzahl von Impfungen ausgewiesen, sofern diese größer Null sind. Für jedes Datum ist angegeben, wie viele Personen, differenziert nach den oben aufgeführten Variablen, geimpft wurden.

Variablenausprägungen

Die Impfdaten enthalten die in der folgenden Tabelle abgebildeten Variablen und deren Ausprägungen:

Variable	Тур	Ausprägung	Beschreibunç
Impfdatum	Datum	JJJJ-MM-TT	Datum der Impfun
BundeslandId_Impfort	Text	01 bis 16 : Bundesland ID 17 : Bundesressorts	Identifikationsnum des Bundeslandes basierend auf den Amtlichen Gemeindeschlüss (AGS). Impfungen des Bundesressor werden separat ausgewiesen, dar Impfstellen des Bundes ohne exa Angabe des Impfc melden
Impfstoff	Text	AstraZeneca : AstraZeneca Moderna : Moderna Comirnaty : BioNTech/Pfizer Janssen : Janssen-Cilag/Johnson & Johnson Novavax : Novavax	Verabreichter Impfstoff
Impfserie	Natürliche Zahl	1 : Erstimpfung2 : Zweitimpfung3 : erste Auffrischimpfung4 : zweite Auffrischimpfung	Angabe zur Erst-, Zweit- oder Auffrischimpfung
Anzahl	Natürliche Zahl	≥1	Anzahl der Impfun in der Impfgruppe

Zweitimpfungen mit dem Janssen-Impfstoff aus dem **KBV-Datenpaket** bzw. Impfungen mit der Abrechnungsziffer für Janssen-Impfstoff aus den **Abrechnungsdaten der KVen**, die vor dem 27.01.2022 erfolgten, werden mit Impfserie 1 ausgewiesen.

COVID-19 Impfdaten auf Ebene der Landkreise

Die aktuellen COVID-19-Impfzahlen auf Ebene der Landkreise sind im Hauptverzeichnis unter "Aktuell_Deutschland_Landkreise_COVID-19_Impfungen.csv" abrufbar und werden täglich überschrieben. Im Archivordner sind die täglichen Impfdaten der Landkreise unter den Dateinamen "JJJJ-MM-TT_Deutschland_Landkreise_COVID-19-Impfungen.csv" abgelegt.

Aktuell_Deutschland_Landkreise_COVID-19_Impfungen.csv Archiv/JJJJ-MM-TT_Deutschland_Landkreise_COVID-19-Impfungen.csv

Die Impfdaten der Landkreise bilden einen tagesaktuellen Stand (08:00 Uhr) aller an das RKI auf Landkreisebene gemeldeten Impfungen in Deutschland ab. Im Dateinamen repräsentiert die Sequenz "JJJJ-MM-TT" das Erstellungsdatum der Datei und gleichzeitig das Datum des enthaltenen Datenstands. "JJJJ" steht dabei für das Jahr, "MM" für den Monat und "TT" für den Tag der Erstellung bzw. des enthaltenen Datenstands. Die "Aktuell_Deutschland_Landkreise_COVID-19-Impfungen.csv" ist identisch mit dem neusten Datenstand des Archivs.

Variablen

Die Impfdaten differenzieren nach verschiedenen Merkmalen einer "Impfgruppe". Unter "Impfgruppe" wird die Zusammenfassung der Merkmalsausprägungen pro Impftag verstanden. Pro Eintrag bzw. Zeile ist eine eineindeutige Impfgruppe für ein Impfdatum abgebildet. Eine Impfgruppe wird grundlegend durch folgende Merkmale charakterisiert (in den Klammern finden sich die Variablen dieser Merkmale):

- Ort der Impfung (Landkreisld_Impfort)
- Angabe des Impfschutzes bzw. der Impfserie (Impfschutz)
- Altersgruppe der Geimpften (Altersgruppe)

Zusätzlich werden folgende Variablen angegeben:

- Datum der Impfung (Impfdatum)
- Anzahl der Impfungen in der Gruppe (Impfungen)

Für jede Impfgruppe wird die tägliche Anzahl von Impfungen ausgewiesen, sofern diese größer vier ist.

Erklärung zur Variable Impfschutz

Da eine Umbewertung der Janssen-Impfungen aus dem **KBV-Datenpaket** mit Altersangabe vor dem 27.01.2022 nicht möglich ist, werden für die Impfungen der niedergelassenen Ärzte in diesen Zeitraum die gemeldeten Impfserien der Impfungen ausgewiesen. Damit unterscheidet sich die Variable Impfschutz von der Variable Impfserie im vorherigen Datensatz. Ab dem 27.01.2022 entspricht die Variable Impschutz der Impfserie im Datensatz COVID-19 Impfdaten auf Ebene der Bundesländer.

Hinweis zur Interpretation der Landkreisangaben

Die Zuordnung der Impfzahlen zu Bundesländern und Landkreisen erfolgt anhand der Postleitzahl (PLZ) der impfenden Stellen (Impfzentren, Betriebe, Betriebsärzt:innen, Apotheken, niedergelassene Ärzteschaft). Nur diese Angabe des Orts der Impfung ist in allen Datenquellen enthalten. Die PLZ der Geimpften ist in den Daten der KBV und der PVS nicht enthalten. Die Wohnort-PLZ ist zwar in den KV-Daten enthalten, diese liegen jedoch nicht bis zum aktuellen Impftag vor. Teilweise ist auch aus den Daten der DIM-Anwendung keine eindeutige Zuordnung zum Wohnort möglich aufgrund unvollständiger Erhebung, unklarer bzw. unbekannter Angaben oder eines Wohnortes außerhalb des Bundesgebietes.

Anhand der Ortsangabe der impfenden Stelle lässt sich keine Impfquote auf Landkreisebene berechnen, da Geimpfte ihren Wohnsitz auch in anderen Landkreisen als dem Landkreis der impfenden Stelle haben können. Zu beachten ist außerdem, dass die Zuordnung von PLZ zu Landkreis nicht immer eindeutig ist. Der Anteil der in einem Landkreis ansässigen Geimpften ist daher auf Grundlage der vorliegenden Daten nicht ermittelbar.

Variablenausprägungen

Die Impfdaten enthalten die in der folgenden Tabelle abgebildeten Variablen und deren Ausprägungen:

Variable	Тур	Ausprägung	Beschreibung
Impfdatum	Datum	JJJJ-MM-TT	Datum der Impfungen
LandkreisId_Impfort	Text	01001 bis 16077 : Landkreis ID 17000 : Bundesressorts u : unbekannt	Identifikationsnummer des Landkreises basierend auf dem Amtlichen Gemeindeschlüssel (AGS). Impfungen des Bundesressorts werden separat ausgewiesen, da die Impfstellen des Bundes ohne exakte Angabe des Impfortes melden.
Altersgruppe	Text	05-11: Altersgruppe 5 bis 11 Jahre 12-17: Altersgruppe 12 bis 17 Jahre 18-59: Altersgruppe 18 bis 59 Jahre 60+: Altersgruppe 60 Jahre und älter	Altersgruppen der in der Impfgruppe enthaltenen Fälle nach Schema der KBV
Impfschutz	Natürliche Zahl	1 : Erstimpfung2 : Grundimmunisierung3 : erste Auffrischimpfung4 : zweite Auffrischimpfung	Angabe zum Impfschutz Grundimmunisierung wird angenommen bei allen gemeldeten Zweitimpfungen sowie Impfungen mit Janssen- Impfstoff bei niedergelassenen Ärzt:innen vor dem 27.01.2022

Variable	Тур	Ausprägung	Beschreibung
Anzahl	Natürliche Zahl	≥5	Anzahl der Impfungen in der Impfgruppe

Ist die Anzahl an Impfungen einer Impfgruppe an einem Tag kleiner als fünf, werden, aus Gründen des Datenschutzes, an diesem Tag keine Impfungen für die Impfgruppe ausgewiesen. Um dennoch einen genauen Überblick über die Gesamtzahl der Impfungen zu ermöglichen, werden Impfgruppen mit weniger als fünf Impfungen zu Impfungen der Folgetage derselben Impfgruppe hinzuaddiert, bis die kumulierte Anzahl der Impfungen an einem Tag den Wert von fünf übersteigt.

Beispiel

Betrachten wir folgende Impfgruppe der über 60ig Jährigen, im Landkreis 01004, die eine Erstimpfung erhalten haben, über drei Tage hinweg:

```
Impfdatum, LandkreisId_Impfort, Altersgruppe, Impfschutz, Anzahl 2021-06-26 ,01004 ,60+ ,1 ,3 2021-06-27 ,01004 ,60+ ,1 ,1 2021-06-28 ,01004 ,60+ ,1 ,12
```

Die Anzahl der Impfungen dieser Impfgruppe ist für die ersten beiden Tage kleiner fünf. In Summe entspricht sie am 27. Juni vier Impfungen, damit ebenfalls kleiner fünf, am 28. Juni sechzehn Impfungen und damit größer fünf. Die Impfungen der ersten beiden Tage werden daher kumuliert mit den Impfungen des dritten Tages ausgewiesen:

```
Impfdatum,LandkreisId_Impfort,Altersgruppe,Impfschutz,Anzahl 2021-06-28 ,01004 ,60+ ,1 ,16
```

COVID-19 Impfquoten

Die aktuellen COVID-19-Impfquoten sind im Hauptverzeichnis unter "Aktuell_Deutschland_Impfquoten_COVID-19.csv" abrufbar und werden täglich überschrieben. Im Archivordner sind die täglichen Impfquoten unter den Dateinamen "JJJJ-MM-TT Deutschland_Impfquoten_COVID-19.csv" abgelegt.

```
Aktuell_Deutschland_Impfquoten_COVID-19.csv
Archiv/JJJJ-MM-TT_Deutschland_Impfquoten_COVID-19.csv
```

Im Dateinamen repräsentiert die Sequenz "JJJJ-MM-TT" das Erstellungsdatum der Datei und gleichzeitig das Datum des enthaltenen Datenstands. "JJJJ" steht dabei für das Jahr, "MM" für den Monat und "TT" für den Tag der Erstellung bzw. des enthaltenen Datenstands. Die "Aktuell_Deutschland_COVID-19-Impfquoten.csv" ist identisch mit dem neusten Datenstand des Archivs.

Die kumulative Zahl der Impfungen umfasst alle Impfungen bis einschließlich des Vortages, die bis zum Erstellungsdatum, 08:00 Uhr, dem RKI gemeldet wurden. Nachmeldungen und Datenkorrekturen aus zurückliegenden Tagen sind in der kumulativen Zahl der Impfungen enthalten.

Die regionale Zuordnung aller durchgeführten Impfungen erfolgt anhand des Ortes der impfenden Stelle und nicht anhand des Wohnortes der geimpften Person. Nur die Angabe des Ortes der impfenden Stellen ist in allen Datenquellen enthalten (siehe Hinweis zu den Datenquellen). Da diese regionalisierten Impfdaten zur

Berechnung einer Impfquote auf die jeweilige Wohnbevölkerung bezogen werden, können dabei rechnerisch auch Anteile von >100% kalkuliert werden.

Alle Impfquoten beziehen sich immer auf die Gesamtzahl der jeweils angegebenen Bevölkerungsgruppe. Die Impfquote "Gesamtbevölkerung" ist der Anteil aller bisher Geimpften in der Gesamtbevölkerung; die Impfquoten nach Alter bilden den Anteil der Geimpften in der jeweiligen Altersgruppe ab. Für die Berechnung der Impfquote wurde der Bevölkerungsstand vom 31.12.2020 zugrunde gelegt (Statistisches Bundesamt (Destatis), 2021, https://www.destatis.de/DE/Themen/Gesellschaft-

Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/bevoelkerung-nichtdeutsch-laender.html). Die zweite Auffrischimpfung wird gegenwärtig von der STIKO nur bestimmten Bevölkerungsgruppen empfohlen. Diese Gruppen werden im Impfquoten-Monitoring nicht spezifisch erfasst. Aus diesem Grund ist es nicht möglich für diese Gruppen spezifische Impfquoten abzubilden und es wird stattdessen auch hier der Bezug zur Wohnbevölkerung nach Altersgruppe hergestellt.

Variablen

Die Tabelle der Impfquoten differenziert grundlegend nach den Merkmalen des Impfstatus und der Altersgruppen.

- Impfstatus (min1, gi, boost1, boost2)
- Altersgruppe (_gesamt, _05bis17, _05bis11, _12bis17, _18bis59, _18plus, _60plus)

Neben den Impfquoten wird zusätzlich die absolute Anzahl der Geimpften - differenziert nach Impfstatus - angegeben. Es werden nur Impfungen einbezogen, die bis zum Vortag des Erstellungsdatums durchgeführt und bis zum Erstellungsdatum der Datei gemeldet wurden.

Variablenausprägungen

Die Tabelle der Impfquoten enthält die in der folgenden Tabelle abgebildeten Variablen und deren Ausprägungen:

Variable	Тур	Ausprägung	Beschreibung
Datum	Datum	JJJJ-MM-TT	Datum, bis zu dem alle durchgeführten und gemeldeten Impfungen berücksichtigt werden
Bundesland	Text	Schleswig-Holstein Thüringen Deutschland	Name des Bundeslandes zuzüglich der Angabe für das gesamte Bundesgebiet

Variable	Тур	Ausprägung	Beschreibung
BundeslandId_Impfort	Text	01 bis 16 : Bundesland ID 17 : Bundesressorts	Identifikationsnummer des Bundeslandes basierend auf dem Amtlichen Gemeindeschlüssel (AGS). Impfungen des Bundesressorts werden separat ausgewiesen, da die Impfstellen des Bundes ohne exakte Angabe des Impfortes melden
Impfungen_gesamt	natürliche Zahl	≥0	Gesamtzahl der aller verabreichten Impfungen
Impfungen_gesamt_min1	natürliche Zahl	≥0	Gesamtzahl mindestens einmal Geimpfter
Impfungen_gesamt_gi	natürliche Zahl	≥0	Gesamtzahl der Grundimmunisierten.
Impfungen_gesamt_boost1	natürliche Zahl	≥0	Gesamtzahl der Personen mit einer ersten Auffrischimpfung
Impfungen_gesamt_boost2	natürliche Zahl	≥0	Gesamtzahl der Personen mit einer zweiten Auffrischimpfung
Impfquote_gesamt_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindesten einmal geimpften Personen
Impfquote_05bis17_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen im Alter von 5 bis 17 Jahren
Impfquote_05bis11_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen im Alter von 5 bis 11 Jahren
Impfquote_12bis17_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen im Alter von 12 bis 17 Jahren
Impfquote_18plus_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindesten einmal geimpften Personen im Alter ab 18 Jahren

Variable	Тур	Ausprägung	Beschreibung
Impfquote_18bis59_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindesten einmal geimpften Personen im Alter von 18 bis 59 Jahren
Impfquote_60plus_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindesten einmal geimpften Personen ab 60 Jahren
Impfquote_gesamt_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen
Impfquote_05bis17_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen im Alter von 5 bis 17 Jahren
Impfquote_05bis11_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen im Alter von 5 bis 11 Jahren
Impfquote_12bis17_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen im Alter von 12 bis 17 Jahren
Impfquote_18plus_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen im Alter ab 18 Jahren
Impfquote_18bis59_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen im Alter von 18 bis 59 Jahren
Impfquote_60plus_gi	rationale Zahl	≥0.0 oder NA	Impfquote der grundimmunisierten Personen ab 60 Jahren
Impfquote_gesamt_boost1	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit erster Auffrischimpfung
Impfquote_12bis17_boost1	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit erster Auffrischimpfung im Alter von 12 bis 17 Jahren

Variable	Тур	Ausprägung	Beschreibung
Impfquote_18plus_boost1	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit erster Auffrischimpfung im Alter ab 18 Jahren
Impfquote_18bis59_boost1	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit erster Auffrischimpfung im Alter von 18 bis 59 Jahren
Impfquote_60plus_boost1	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit erster Auffrischimpfung ab 60 Jahren
Impfquote_gesamt_boost2	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit zweiter Auffrischimpfung
Impfquote_12bis17_boost2	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit zweiter Auffrischimpfung im Alter von 12 bis 17 Jahren
Impfquote_18plus_boost2	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit zweiter Auffrischimpfung im Alter ab 18 Jahren
Impfquote_18bis59_boost2	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit zweiter Auffrischimpfung im Alter von 18 bis 59 Jahren
Impfquote_60plus_boost2	rationale Zahl	≥0.0 oder NA	Impfquote der Personen mit zweiter Auffrischimpfung ab 60 Jahren

Erklärung zur Variablenausprägung "Bundesressorts"

Umfasst Impfungen, die aus dem Impfkontingent des Bundes gemäß Coronavirus-Impfverordnung an Angehörige des Bundes verabreicht wurden. Eine Impfquote kann aufgrund einer fehlenden Nennerpopulation nicht berechnet werden, die Impfungen gehen allerdings in die Berechnung der Impfquote für Gesamtdeutschland mit ein.

Die entsprechenden Variablen sind deshalb um den Wert "NA" in ihrer Ausprägung ergänzt, was bei der Verarbeitung dieser Variablen berücksichtigt werden sollte.

Erklärung zum Merkmal Impfstatus

Die Gesamtzahl mindestens einmal Geimpfter (min1) umfasst alle Meldungen, für die in den Daten des DIM die Impfserie 1 angegeben ist. Als Impfserie 1 gelten Erstimpfungen mit den in Deutschland zugelassenen und verfügbaren Impfstoffen. Nicht enthalten sind Impfungen von Personen, die erstmalig in Deutschland geimpft wurden, zuvor aber von einer SARS-CoV-2 Infektion genesen sind oder bereits eine Impfung im Ausland erhalten hatten. Diese gelten direkt als grundimmunisiert.

Darüber hinaus gelten alle Personen als grundimmunisiert (gi), für die in den Daten des DIM die Impfserie 2 angegeben ist. Darunter zählen Zweitimpfungen in Deutschland und Impfungen nach Genesung.

Als Personen mit erster bzw. zweiter Auffrischimpfung (boost1, boost2) gelten Personen, die eine oder zwei weitere Impfungen nach abgeschlossener Grundimmunisierung erhalten haben. Diese werden mit Impfserie 3 bzw. 4 gemeldet. Auch für die Auffrischungsimpfungen gilt, dass diesen eine Impfung im Ausland vorrangegangen sein kann und in den Daten somit keine vorhergehende Impfserie 1 und Impfserie 2 gemeldet wurde.

Ziel des Impfquotenmonitorings ist es, eine Aussage zum Impfschutzniveau auf Bevölkerungsebene zu treffen. Die Datenerhebung folgt dieser Perspektive. Da vom DIM keine Daten zum Genesungsstatus oder bestehenden Impfungen im Ausland erhoben werden, ist in der Datenauswertung keine Aussage oder Berücksichtigung diesbezüglich möglich.

Unschärfen der Impfquoten

Da alle bis zum 26.01.22 von den Vertragsärzt:innen gemeldeten Janssen-Impfungen ausschließlich als Zweit bzw. erste Auffrischimpfungen gemeldet wurden aber unklar ist, wie viele tatsächliche Erst- oder Zweitimpfungen darunter waren, wird die Impfquote der Grundimmunisierten systematisch zu hoch ausgewiesen.

Da nicht rekonstruierbar ist, wie in den Meldeportalen von PVS und KBV die empfohlenen Folgeimpfungen nach einmaliger Janssen-Impfung gemeldet wurden, bestehen weitere Unschärfen bei den Quoten der Grundimmunisierung und der ersten Auffrischimpfungen.

Gesonderte Ausweisung der Kinderimpfungen bei 5-11-Jährigen

Seit dem 21.12.2021 werden Kinderimpfungen bei 5-11-Jährigen gesondert ausgewiesen. Bis einschließlich 20.12.2021 wurden alle Kinderimpfungen der Altersgruppe 12-17 Jahre zugewiesen.

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-19-Impfungen 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: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.