

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)

FG 31 | Infektionsepidemiologische Fach-IT und Anwendungsentwicklung

[Lei Mao](#) (Projektleitung)

[Hermann Claus](#) (Technische Leitung)

Constantin Fischer (Datenmanagement)

Sven Schröder (Datenmanagement)

MF 4 | Forschungsdatenmanagement

[Hannes Wuensche](#) (Datenkuration)

Robert Koch-Institut (2021): COVID-19-Impfungen in Deutschland, Berlin: Zenodo.

DOI:[10.5281/zenodo.5126652](https://doi.org/10.5281/zenodo.5126652).

Der Datensatz "COVID-19-Impfungen in Deutschland" ist lizenziert unter der [Creative Commons Namensnennung 4.0 International Public License](#) | [CC-BY 4.0 International](#)

Informationen zum Datensatz und Entstehungskontext

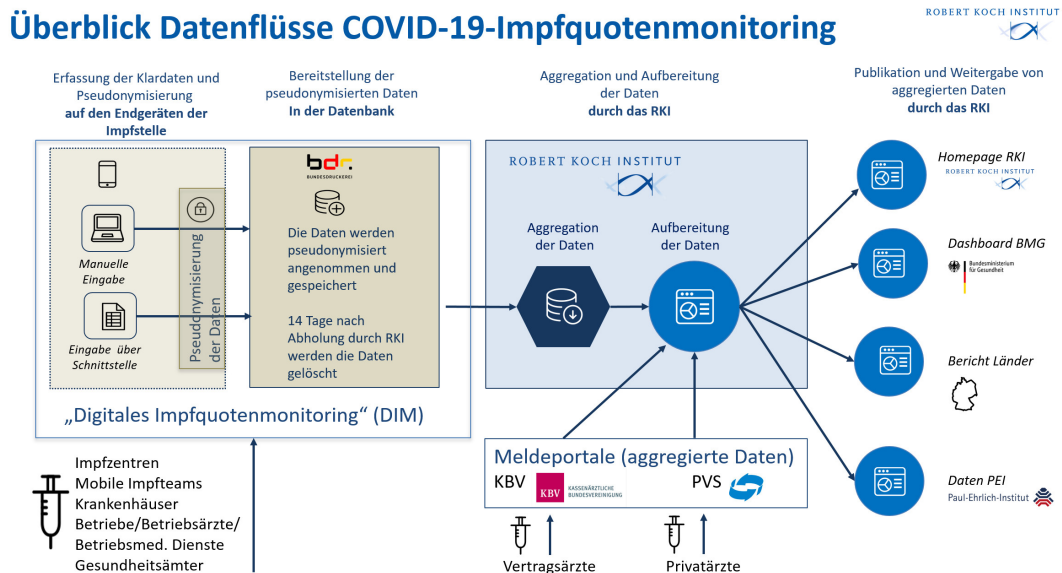
Die COVID-19-Impfung kann einen Wendepunkt in der Kontrolle der COVID-19-Pandemie darstellen und erfährt daher hohes Maß an öffentlicher Aufmerksamkeit. Einführung und Umsetzung der COVID-19-Impfung gehen mit besonderen Herausforderungen einher, die bei der Impfdatenerfassung zu berücksichtigen sind.

In diesem Kontext ist es Ziel des Projekts "Digitales Impfquoten-Monitoring" (DIM), tagesaktuell und bundesweit die Impfquote zu erfassen und folgend aufbereitet darzustellen, um zeitnah den Verlauf der COVID-19-Impfkampagne zu analysieren, bei Bedarf nachzusteuern und logistische bzw. organisatorische Konsequenzen zu ziehen.

Ein Hauptbestandteil des DIM-Projekts ist die DIM-Anwendung, eine Web-basierte Softwarelösung zur Erfassung der Impfquote der Bevölkerung gegen SARS-CoV2 auf Grundlage der [Coronavirus-Impfverordnung \(CoronaImpfV\)](#).

Die DIM-Anwendung besteht aus Eingabemasken/Dialogen (Frontend) und dahinter liegenden Datenbanksystem (Backend). Autorisiertes Personal der stationären Impfzentren und mobile Impfteams

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 an in das Datenbanksystem übertragen werden. Die nachfolgende Darstellung zeigt schematisch den Datenfluss von Erfassung der Daten in den Impfzentren bis hin zur Weitergabe der aggregierten Daten durch das RKI.



Projektbeteiligte und Rollenbesetzung

An dem Projekt "Digitales Impfquoten-Monitoring" sind verschiedene Stakeholder beteiligt. Projektauftraggeber ist das [Bundesministerium für Gesundheit \(BMG\)](#). Das [Robert Koch-Institut](#) übernimmt die Projektleitung. Die [Bundesdruckerei \(BdR\)](#) ist zuständig für die technische Entwicklung und den Betrieb der Web-Anwendung. [Accenture GmbH \(ACN\)](#) unterstützt das RKI bei der Konzeption, der Planung, Dokumentation und der Rolloutkoordination. Projektzeitraum ist vom 01.11.2020 bis 30.11.2022.

Projektleitung

- Anette Siedler | [Fachgebiet 33](#)
- Lei Mao und Hermann Claus | [Fachgebiet 31](#)

Die zugrundeliegenden Daten werden über die Bundesdruckerei an das Robert Koch-Institut (RKI) übermittelt. Die Verarbeitung und Aufbereitung der vorliegenden Rohdaten erfolgt durch das [Fachgebiet 31 | Infektionsepidemiologische Fach-IT und Anwendungsentwicklung](#) des RKI. Die epidemiologische Datenanalyse erfolgt durch das [Fachgebiet 33 | Impfprävention](#) des RKI. 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. Diese Daten werden der Disziplin der Infektionsepidemiologie zugeordnet. Die hier veröffentlichten Impfdaten sind aggregierte Daten aus drei Datenquellen. Neben dem Meldeweg über die DIM-Anwendung werden zur Erfassung der Impfquote auch Impfdaten aus Arztpraxen über die Kassenärztliche Bundesvereinigung (KBV) und den Privatärztlicher Bundesverband (PBV) erfasst. Insgesamt werden im Datensatz "COVID-19-Impfungen in Deutschland" somit verschiedenste Datenquellen in einem Datensatz zusammengeführt und ständig erweitert:

- Die DIM-Daten enthalten Angaben der Impfzentren, mobilen Impfteams, Krankenhäuser und der Betriebsärzt:innen, die über die DIM-Webanwendung übermittelt werden.

Aus den Impfzentren, mobilen Impfteams und Krankenhäusern der Bundesländer werden seit dem 27.12.2020, von Betriebsärzt:innen und betriebsmedizinischen Diensten seit dem 07.06.2021, täglich pseudonymisierte Daten mit der Web-Anwendung "Digitales Impfquotenmonitoring" übermittelt.

- Der täglich aggregierte Kerndatensatz der impfenden Vertragsärzt:innen über die [Kassenärztliche Bundesvereinigung \(KBV\)](#).

Von den Vertragsärzt:innen werden täglich aggregierte Daten an die KBV übermittelt und dort in zwei separaten Datenpaketen dem RKI bereitgestellt: Anzahl durchgeführter Impfungen je Praxis und Impftag aufgeschlüsselt nach Impfstoff und Impfstoffdosis (verfügbar seit Impftag 10.03.2021) sowie aufgeschlüsselt nach Altersgruppe <18, 18-59 und 60+ Jahre und Impfstoffdosis (verfügbar seit Impftag 26.03.2021).

- Der täglich aggregierte Kerndatensatz der impfenden Privatärzt:innen über die [Privatärztlicher Bundesverband \(PBV\)](#).

Von den Privatärzt:innen werden täglich aggregierte Daten an die PVS (Privatärztliche Verrechnungsstelle als technischer Dienstleister des PBV) übermittelt und dort in einem Datenpaket dem RKI bereitgestellt: 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).

Aufbereitung der übermittelten Daten

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

- Grundlegende Validitätsprüfung der übermittelten Datensätze
- Einheitliche Benennung der Impfstoffe
- Zuweisung der BundeslandID oder LandkreisID des Imports
- Generelle Ausweisung von Impfungen mit dem Impfstoff Janssen als Erstimpfung (Impfserie = 1) bzw. Auffrischungsimpfung (Impfserie = 3)
- Ausschluss von Datensätzen ohne Angabe der Impfserie oder des Impfzentrums (nur DIM)
- Erkennung von Sammelpatienten:innen (nur DIM)
- Auswahl der aktuellsten Datensätze je Pseudonym und Impfdatum (nur DIM)
- Filterung auf Impfdaten im Zeitraum vom 27.12.2020 bis zum vorhergehenden Tag

Zu beachten ist, dass Zweitimpfungen erst ab dem 15.01.2021 und Auffrischungsimpfungen erst ab 01.06.2021 als plausibel erachtet werden. Entsprechende Impfungen mit einem früheren Datum sind am ehesten auf Eingabefehler zurückzuführen, werden jedoch in den aggregierten Daten mit ausgewiesen.

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.

Korrektur 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 Postleitzahl der Person sowie dem Impfzentrum vor, ist davon auszugehen, dass es sich um Sammelpatient:innen, d.h. verschiedene Personen, handelt. Die Annahme wird auch dann getroffen, wenn Unterschiede in der Postleitzahl der Person oder dem angegebenen Impfzentrum 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 werden folgenden Regeln auf Einträge mit gleichem Personen-Pseudonym/erkannten Sammelpatient:innen angewendet. Eine Auffrischungsimpfung ist erlaubt, wenn sie nach dem 01.06.2021 erfolgte. Der Hinweis "aktuellste" bezieht sich dabei auf das Datum der Datenerfassung, nicht das Impfdatum:

Schritt	Kriterium	Aktion
1	Späteste Erstimpfung mit Janssen	ja : behalte nur die Erstimpfung zum spätesten Impfdatum mit Janssen sowie die erste spätere, erlaubte Auffrischungsimpfung, wenn diese mit Janssen erfolgte. nein : weiter mit Schritt 2
2	Weniger als 14 Tagen max. Abstand im Impfdatum	Ja : Die Einträge werden als Korrekturen voneinander angesehen, daher wird nur der aktuellste Eintrag behalten. Nein : weiter mit Schritt 3
3	Genau eine Erst- vor der Zweitimpfung	Ja : valide Einträge, keine weitere Aktion Nein: weiter mit Schritt 4
4	Genau eine Zweit- vor der Erstimpfung	Ja : Impfsérie vertauscht, tausche die Impfsérie Nein: weiter mit Schritt 5
5	Genau eine Erst- vor der erlaubten Auffrischungsimpfung	Ja : valide Einträge, keine weitere Aktion Nein: weiter mit Schritt 6
6	Genau eine Erst- vor der nicht erlaubten Auffrischungsimpfung	Ja : Ist die Auffrischungsimpfung nicht mit Janssen, dann ändere die Auffrischungsimpfung zur Zweitimpfung ab, sonst verwirfe die Auffrischungsimpfung Nein: weiter mit Schritt 7
7	Genau eine Auffrischungs- vor der Erstimpfung	Ja : Ändere die Auffrischungsimpfung zur Erstimpfung ab. Ist die Erstimpfung mind. 120 Tage später erfolgt und erlaubt im Sinne einer Auffrischungsimpfung, dann ändere die Erstimpfung zur Auffrischungsimpfung ab, sonst zur Zweitimpfung. Nein: weiter mit Schritt 8
8	Genau eine Zweit- vor der erlaubten Auffrischungsimpfung	Ja : valide Einträge, keine weitere Aktion Nein: weiter mit Schritt 9

Schritt	Kriterium	Aktion
9	Genau eine Zweit- vor der nicht erlaubten Auffrischungsimpfung	Ja : Ändere die Zweitimpfung zur Erstimpfung ab. Ist die Auffrischungsimpfung nicht mit Janssen, dann ändere die Auffrischungsimpfung zur Zweitimpfung ab, sonst verwirfe die Auffrischungsimpfung Nein: weiter mit Schritt 10
10	Genau eine Auffrischungs- vor der Zweitimpfung	Ja : Ändere die Auffrischungsimpfung zur Erstimpfung ab. Nein: weiter mit Schritt 11
11	Sonst	Das früheste Impfeignis wird zur Erstimpfung. Das erste folgende Impfeignis mit Impfstoff ungleich Janssen und mind. 14 Tagen Abstand wird zur Zweitimpfung. Die erste folgende, erlaubte Auffrischungsimpfung wird zur Auffrischungsimpfung.

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

Aufbau und Inhalt des 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 Datei enthalten. Der verwendete Zeichensatz der .csv Datei 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 (8:30 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. Pro Eintrag bzw. Zeile ist eine eindeutige Impfgruppe 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 Variablen angegeben:

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

Eine Impfgruppe nimmt eine eindeutige Ausprägung hinsichtlich der Anzahl der Impfungen in einem Bundesland, des Impfstoffes und der Impfserie an. Für jede Impfgruppe wird die tägliche Anzahl neuer 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	Typ	Ausprägung	Beschreibung
Impfdatum	Datum	JJJJ-MM-TT	Datum der Impfungen
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
Impfstoff	Text	AstraZeneca: AstraZeneca Moderna: Moderna Comirnaty: BioNTech/Pfizer Janssen: Janssen-Cilag/Johnson & Johnson	Verabreichter Impfstoff
Impfserie	Natürliche Zahl	1: Erstimpfung 2: Zweitimpfung 3: Auffrischungsimpfung	Angabe zur Erst-, Zweit- oder Auffrischungsimpfung

Variable	Typ	Ausprägung	Beschreibung
Anzahl	Natürliche Zahl	≥ 1	Anzahl der Impfungen in der Impfgruppe

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 (8:30 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. Pro Eintrag bzw. Zeile ist eine eindeutige Impfgruppe abgebildet. Eine Impfgruppe wird grundlegend durch folgende Merkmale charakterisiert (in den Klammern finden sich die Variablen dieser Merkmale):

- Ort der Impfung (LandkreisId_Impfort)
- Angabe zur Vollständigkeit/Unvollständigkeit des Impfschutzes (Impfschutz)
- Altersgruppe der Geimpften (Altersgruppe)

Zusätzlich werden folgende Variablen angegeben:

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

Eine Impfgruppe nimmt eine eindeutige Ausprägung hinsichtlich der Anzahl der Impfungen eines Landkreises, der Altersgruppe der Geimpften und der Ausprägung des Impfschutzes an. Für jede Impfgruppe wird die tägliche Anzahl neuer Impfungen ausgewiesen, sofern diese größer vier ist.

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, niedergelassene Ärzteschaft). Nur diese Angabe des Orts der Impfung ist in allen Datenquellen enthalten. Die PLZ der Geimpften ist in den Daten der KBV nicht enthalten, teilweise fehlt sie, aufgrund unvollständiger Übermittlung, in den Daten des DIM.

Anhand der PLZ 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. Der Anteil der in einem Landkreis ansässigen Geimpften ist daher auf Grundlage der vorliegenden Daten nicht ermittelbar.

Hinweis zur Berechnung von Impfquoten auf Basis der Landkreisdaten

Auf Basis der COVID-19 Impfdaten der Landkreise lässt sich keine Berechnung der Impfquoten, wie sie in den Daten der Pressetabelle des RKI vorliegen, durchführen.

Hintergrund ist, dass über die Variable "Impfschutz", keine Aussage zur Anzahl der mindestens einmal Geimpften getroffen werden kann. Als "mindestens einmal geimpft" gelten Personen, die eine Erstimpfung mit den Impfstoffen von AstraZeneca, BioNTech oder Moderna oder eine Einzelimpfung mit Janssen erhalten haben. Personen, die mit Janssen geimpft wurden, gelten daher als "mindestens einmal geimpft", werden aber in der Tabelle der Landkreise ausschließlich in der Gruppe der vollständig Geimpften ausgewiesen (Impfschutz = 2). Eine Ableitung mindestens einmal geimpfter Personen ist nicht möglich, da aus den Daten nicht

ermittelbar ist, wie viele Impfungen mit vollständigem Impfschutz auf Impfungen mit Janssen entfallen. Die Impfquoten werden daher separat bereitgestellt, siehe [COVID-19 Impfquoten](#)

Variablenausprägungen

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

Variable	Typ	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	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: Unvollständiger Impfschutz 2: Vollständiger Impfschutz 3: Aufgefrischter Impfschutz	Angabe zum Impfschutz Vollständiger Impfschutz besteht bei zweifacher Impfung, Impfung mit Janssen und einfach Geimpften mit überstandener SARS-CoV-2 Infektion
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_COVID-19_Impfquoten.csv" abrufbar und werden täglich überschrieben. Im Archivordner sind die täglichen Impfquoten unter den Dateinamen "JJJJ-MM-TT_Deutschland_COVID-19-Impfquoten.csv" abgelegt.

```
Aktuell_Deutschland_COVID-19-Impfquoten.csv
Archiv/JJJJ-MM-TT_Deutschland_COVID-19-Impfquoten.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.

Variablen

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

- Impfstatus (_min1, _voll, _boost)
- Altersgruppe (_gesamt, _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	Typ	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
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 Impfungen. Die Impfungen mit Janssen sind sowohl in der Gruppe der mindestens einmal Geimpften als auch in der Gruppe der vollständig Geimpften enthalten. Sie werden für die Gesamtzahl der verabreichten Impfungen jedoch nur einmal gezählt.
Impfungen_gesamt_min1	natürliche Zahl	≥0	Gesamtzahl mindestens einmal Geimpfter. Diese umfasst alle Personen, die Erstimpfungen mit den Impfstoffen von BioNTech, Moderna oder AstraZeneca oder eine Impfung mit dem Impfstoff Janssen erhalten haben.
Impfungen_gesamt_voll	natürliche Zahl	≥0	Gesamtzahl der vollständig Geimpften. Diese umfasst alle Personen, die Zweitimpfungen mit BioNTech, Moderna oder AstraZeneca oder eine Impfung mit Janssen erhalten haben.
Impfungen_gesamt_boost	natürliche Zahl	≥0	Gesamtzahl der Personen mit einer Auffrischungsimpfung
Impfquote_gesamt_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindesten einmal geimpften Personen
Impfquote_12bis17_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen im Alter von 12 bis 17 Jahren

Variable	Typ	Ausprägung	Beschreibung
Impfquote_18plus_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen im Alter ab 18 Jahren
Impfquote_18bis59_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen im Alter von 18 bis 59 Jahren
Impfquote_60plus_min1	rationale Zahl	≥0.0 oder NA	Impfquote der mindestens einmal geimpften Personen ab 60 Jahren
Impfquote_gesamt_voll	rationale Zahl	≥0.0 oder NA	Impfquote der vollständig geimpften Personen
Impfquote_12bis17_voll	rationale Zahl	≥0.0 oder NA	Impfquote der vollständig geimpften Personen im Alter von 12 bis 17 Jahren
Impfquote_18plus_voll	rationale Zahl	≥0.0 oder NA	Impfquote der vollständig geimpften Personen im Alter ab 18 Jahren
Impfquote_18bis59_voll	rationale Zahl	≥0.0 oder NA	Impfquote der vollständig geimpften Personen im Alter von 18 bis 59 Jahren
Impfquote_60plus_voll	rationale Zahl	≥0.0 oder NA	Impfquote der vollständig geimpften Personen ab 60 Jahren

Für die Bundesressorts können keine Impfquoten ausgewiesen werden. Die entsprechenden Variablen sind deshalb um den Wert "NA" in ihrer Ausprägung ergänzt, was bei der Verarbeitung dieser Variablen berücksichtigt werden sollte.

Hinweis bezüglich Impfungen mit Impfstoff Janssen

Die Gesamtzahl mindestens einmal Geimpfter umfasst alle Personen, die Erstimpfungen mit den Impfstoffen von BioNTech, Moderna oder AstraZeneca oder eine Impfung mit dem Impfstoff Janssen erhalten haben. Als vollständig geimpft gelten alle Personen, die Zweitimpfungen mit BioNTech, Moderna oder AstraZeneca oder eine Impfung mit Janssen erhalten haben. Die Impfungen mit Janssen sind daher sowohl in der Gruppe "mindestens einmal geimpft" als auch in der Gruppe "vollständig geimpft" enthalten. Sie werden für die Gesamtzahl der verabreichten Impfungen jedoch nur einmal gezählt.

Unterschätzung der Impfquoten nach Alter, mindestens einmal Geimpfter

Aus dem vertragsärztlichen Bereich, der rund 40% aller Impfungen ausmacht, werden dem RKI keine Impfdaten aufgeschlüsselt nach Impfstoff und Alter übermittelt. Das RKI erhält von der KBV zwei Datenpakete, das eine aufgeschlüsselt nach Impfstoff und Impfsérie und das andere aufgeschlüsselt nach Altersgruppe und Impfsérie (siehe: [Zusammensetzung der Datenquellen](#)).

Da die KBV sämtliche Janssen-Impfungen als Impfdosis = 2 übermittelt, können die Janssen-Impfungen bei den vollständig Geimpften aller Altersgruppen mitgezählt werden (da in dem Datenpaket nach Alter bei den jeweiligen Altersgruppen für Impfsérie=2 enthalten).

In den Impfquoten der mindestens einmal geimpften Erwachsenen "Impfquote_18plus_min1"), werden alle Janssen-Impfungen der Vertragsärzt:innen unter der Annahme zusammengefasst, dass mit diesem Impfstoff ausschließlich Personen ab 18 Jahre geimpft wurden (entsprechend der Zulassung). Dem RKI ist jedoch unbekannt, wie sich die Janssen-Impfungen der KBV auf die Altersgruppen 18-59 Jahre und 60+ Jahre verteilen. Daher können sie den mindestens einmal Geimpften der beiden Altersgruppen nicht zugewiesen werden. Dadurch werden in beiden Altersgruppen die Impfquoten der mindestens einmal Geimpften systematisch zu niedrig ausgewiesen. Es kommt hinzu, dass generell (in allen 3 Datenquellen) Genesene als "Vollständig Geimpft" übermittelt werden. In diesen Fällen werden ebenfalls keine Erstimpfungen angegeben. Das zusammengefasst führt dazu, dass in einigen Bundesländer die Impfquote der vollständig Geimpften höher als die der mindestens einmal Geimpften ist.

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.

Die empfohlene Zitierweise ist:

Robert Koch-Institut (2021): COVID-19-Impfungen in Deutschland, Berlin: Zenodo. DOI:[10.5281/zenodo.5126652](https://doi.org/10.5281/zenodo.5126652).