

# Hochfrequente Mental Health Surveillance

---

*Hochfrequente Beobachtung der psychischen Gesundheit der erwachsenen Bevölkerung in Deutschland*

Robert Koch-Institut | RKI

Nordufer 20  
13353 Berlin

Stephan Junker, Stefan Damerow, Lena Walther und Elvira Mauz  
Fachgebiet 26 | Psychische Gesundheit

---

## Zitieren

Junker, Damerow, Walther und Mauz (2023): Hochfrequente Mental Health Surveillance, Berlin: Zenodo. DOI: [10.5281/zenodo.8238471](https://doi.org/10.5281/zenodo.8238471)

## Einleitung

---

Im Rahmen der Mental Health Surveillance (MHS) am Robert Koch-Institut (RKI) werden für eine Auswahl an Indikatoren der psychischen Gesundheit von Erwachsenen basierend auf Surveydaten Zeitreihen bestehend aus gleitenden Drei-Monats-Schätzern und Glättungskurven berechnet. Dadurch sollen Entwicklungen in der psychischen Gesundheit der erwachsenen Bevölkerung in Deutschland mit möglichst geringem Zeitverzug beobachtet und insbesondere negative Entwicklungen frühzeitig erkannt werden. Diese hochfrequente Surveillance wurde ursprünglich vor dem Hintergrund neuer Informationsbedarfe zur Entwicklung der psychischen Gesundheit der Bevölkerung in der COVID-19-Pandemie entwickelt.

## Datensatz und Entstehungskontext

---

Im Datensatz „Hochfrequente Mental Health Surveillance“ werden aktuelle Zeitreihen bestehend aus gleitenden Drei-Monats-Schätzern samt Konfidenzintervallen sowie wöchentlichen Schätzern für die Glättungskurven zu den folgenden Indikatoren der psychischen Gesundheit Erwachsener bereitgestellt:

- depressive Symptome
- Angstsymptome
- selbsteingeschätzte psychische Gesundheit

Der Datensatz umfasst je Indikator sowohl Schätzer für die kategoriale Auswertung der entsprechenden Messinstrumente als auch Mittelwerte von Summenwerten. Neben den Schätzern für die erwachsene Gesamtbevölkerung sind auch Schätzer für Bevölkerungsgruppen stratifiziert (aufgeteilt) nach Geschlecht (wie bei Geburt zugewiesen), Alter und Bildung enthalten. Für die Bevölkerungsgruppen werden standardisierte (nach den jeweiligen anderen beiden Charakteristika aus Geschlecht, Alter und Bildung) und nicht standardisierte Schätzer angegeben.

Die 2019 bzw. 2021 beginnenden Zeitreihen werden in der Regel monatlich aktualisiert, sobald neue Daten aus der [GEDA-Studie](#) (siehe Datenerhebung) verfügbar sind.

## Administrative und organisatorische Angaben

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

## Datenerhebung und Datenauswertung

---

Die Datengrundlage der Zeitreihen sind vor allem verschiedene Erhebungswellen der im Rahmen des Gesundheitsmonitorings durchgeführten Studie „[Gesundheit in Deutschland aktuell \(GEDA\)](#)“: [GEDA/EHIS 2019-2020](#), [GEDA 2020](#), [GEDA 2021](#), [GEDA 2022](#) und [GEDA 2023](#). Eine Erhebungslücke in 2021 konnte mit Daten des „[COVID-19 Impfquoten-Monitoring in Deutschland \(COVIMO\)](#)“ geschlossen werden. Alle Studien wurden durch das Bundesministerium für Gesundheit beauftragt und für die Auswertungen zusammengeführt. Diese repräsentativ angelegten telefonischen Gesundheitssurveys umfassen monatliche Erhebungen in Zufallsstichproben von zunächst jeweils ca. 1.000 in Deutschland lebenden Erwachsenen (2019-2021), dann von ca. 3.000 Erwachsenen (2022-2023) und aktuell von ca. 1.000 bzw. für depressive Symptome 2.000 Erwachsenen (ab Mitte April 2023). Die Stichprobenziehung erfolgte in diesen Studien über ein sogenanntes Dual-Frame-Verfahren, das eine Mobilfunk und eine Festnetzgesamtheit als Auswahlgesamtheiten nutzt ([Allen et al., 2021](#)).

Zur Durchführung der telefonischen Befragungen arbeitet das RKI mit einem externen Markt- und Sozialforschungsinstitut (USUMA GmbH) zusammen, welches im Auftrag des Epidemiologischen Daten- und Befragungszentrums der Abteilung 2 des Robert Koch-Instituts die Daten erhebt und unter Einhaltung der Datenschutz- und Sicherheitsbestimmungen anonymisiert zur Verfügung stellt. Bevor die Daten für Auswertungen bereitgestellt werden, findet eine standardisierte Datenaufbereitung inklusive Datenqualitätssicherung statt. So werden z.B. unplausible Angaben identifiziert, Fälle bereinigt oder Variablen generiert ([Allen et al., 2021](#)). Nach der Bereitstellung der Daten für die MHS erfolgt eine inhaltliche Datenprüfung bezüglich der Angaben zur psychischen Gesundheit und es werden Gewichtungsfaktoren berechnet, die unterschiedliche Teilnahmebereitschaften in verschiedenen Bevölkerungsgruppen sowie die verschiedenen Auswahlwahrscheinlichkeiten der befragten Personen berücksichtigen. Nähere Studieninformationen sind abrufbar unter:

“Gesundheit in Deutschland aktuell - European Health Interview Survey” | GEDA-EHIS 2019-2020, GEDA 2021, GEDA 2022, GEDA 2023:

[https://www.rki.de/DE/Content/Gesundheitsmonitoring/Studien/Geda/Geda\\_node.html](https://www.rki.de/DE/Content/Gesundheitsmonitoring/Studien/Geda/Geda_node.html)

COVIMO-Studie: Impfverhalten, Impfbereitschaft und -akzeptanz in Deutschland:

[https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Projekte\\_RKI/covimo\\_studie\\_Ergebnisse.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/covimo_studie_Ergebnisse.html)

## Untersuchungszeitraum

Die Untersuchungszeiträume unterscheiden sich für die verschiedenen Indikatoren. Dies ergibt sich aus Unterschieden zwischen den Datengrundlagen mit Blick auf die eingesetzten Messinstrumente. Depressive Symptome werden seit April 2019 beobachtet; Angstsymptome und die selbsteingeschätzte psychische Gesundheit seit März 2021.

## Grundgesamtheit

Die Grundgesamtheit der GEDA- und COVIMO-Erhebungen umfasst die deutschsprachige, in Privathaushalten lebende Bevölkerung ab 15 Jahren (GEDA-EHIS 2019/2020), 16 Jahren (GEDA 2021, 2022) bzw. 18 Jahren (GEDA 2023, COVIMO), deren üblicher Aufenthaltsort zum Zeitpunkt der Datenerhebung in Deutschland liegt. Hierbei sind sowohl Einzelpersonen- als auch Mehrpersonenhaushalte inbegriffen, die eigenständig wirtschaften und sich selbstständig mit für den Lebensunterhalt notwendigen Dingen versorgen. Diese Definition umfasst keine kollektiven Haushalte wie Krankenhäuser, Pflege- oder Wohnheime, Gefängnisse, Kasernen, religiöse Einrichtungen, Pensionen oder Wohnheime. "Üblicher Aufenthaltsort" beschreibt den Ort, an dem eine Person normalerweise lebt oder ihren Lebensmittelpunkt sieht – ungeachtet vorübergehender Abwesenheit zu Zwecken der Erholung, der beruflichen Tätigkeit, der medizinischen Behandlung oder ähnlichem.

Die Auswertungen für den vorliegenden Datensatz wurden mit GEDA- und COVIMO-Daten von Erwachsenen (ab 18 Jahren) durchgeführt, weil die Entwicklung der psychischen Gesundheit Erwachsener im Fokus stehen. Ab 2022 sind die untersuchten Indikatoren außerdem nicht mehr bei unter 18-Jährigen abgefragt worden.

## Erhebungsinstrumente

Der vorliegende Datensatz enthält aggregierte Daten aus den Studien GEDA/EHIS 2019-2020, GEDA 2020, COVIMO, GEDA 2021, GEDA 2022 und GEDA 2023 zu drei Indikatoren:

### Indikator „selbsteingeschätzte psychische Gesundheit“

Der Indikator „selbsteingeschätzte psychische Gesundheit“ beinhaltet eine Selbsteinschätzung der eigenen psychischen Gesundheit im Allgemeinen. Die Messung erfolgt über ein etabliertes Einzelitem (Ahmad et al., 2014):

„Wie würden Sie Ihren psychischen Gesundheitszustand im Allgemeinen beschreiben?“

Antwortoptionen: ausgezeichnet = 5, sehr gut = 4, gut = 3, weniger gut = 2, schlecht = 1

Der Datensatz enthält Schätzer zum Bevölkerungsmittelwert sowie zum Anteil mit einem Wert  $\geq 4$ , also dem Anteil in Deutschland lebender Erwachsener mit einer sehr guten bis ausgezeichneten selbsteingeschätzten psychischen Gesundheit.

### Indikator "depressive Symptome"

Der Indikator "depressive Symptome" wird mit dem validierten Messinstrument „Patient Health Questionnaire-2“ (PHQ-2) erfasst (Löwe et al., 2005). Der PHQ-2 fragt:

„Wie oft fühlten Sie sich im Verlauf der letzten 2 Wochen durch die folgenden Beschwerden beeinträchtigt“:

1. „Wenig Interesse oder Freude an Ihren Tätigkeiten“
2. „Niedergeschlagenheit, Schwermut oder Hoffnungslosigkeit“

Antworten und Punktzahlen jeweils: überhaupt nicht = 0, an einzelnen Tagen = 1, an mehr als der Hälfte der Tage = 2, beinahe jeden Tag = 3

Der Summenwert des PHQ-2 beträgt 0 (min) bis 6 (max) und weist auf keine bis sehr starke Belastung durch depressive Symptome hin. Werte  $\geq 3$  deuten darauf hin, dass eine auffällige Belastung durch depressive Symptome vorliegt. Der Datensatz enthält Schätzer zum Bevölkerungsmittelwert sowie zum Bevölkerungsanteil im auffälligen Wertebereich.

### Indikator "Angstsymptome"

Der Indikator "Angstsymptome" wird mit dem validierten Messinstrument „Generalized Anxiety Disorder-2“ (GAD-2) erfasst (Kroenke et al., 2007). Der GAD-2 fragt:

„Wie oft fühlten Sie sich im Verlauf der letzten 2 Wochen durch die folgenden Beschwerden beeinträchtigt“:

1. „Nervosität, Ängstlichkeit oder Anspannung“
2. „Nicht in der Lage sein, Sorgen zu stoppen oder zu kontrollieren“

Antworten und Punktzahlen jeweils: überhaupt nicht = 0, an einzelnen Tagen = 1, an mehr als der Hälfte der Tage = 2, beinahe jeden Tag = 3

Der Summenwert des GAD-2 beträgt 0 (min) bis 6 (max) und weist auf keine bis sehr starke Belastung durch Angstsymptome hin. Werte  $\geq 3$  deuten darauf hin, dass eine auffällige Belastung durch Angstsymptome vorliegt. Der Datensatz enthält Schätzer zum Bevölkerungsmittelwert sowie zum Bevölkerungsanteil im auffälligen Wertebereich.

## Datenauswertung

Um die Entwicklung mehrerer Indikatoren der psychischen Gesundheit der in Deutschland lebenden erwachsenen Bevölkerung über die Zeit zu untersuchen, werden Zeitreihen bestehend aus monatlichen geschätzten Werten samt Konfidenzintervallen sowie Glättungskurven berechnet.

### **Gleitende Drei-Monats-Schätzer**

Bei den monatlichen Schätzern handelt es sich um gleitende Drei-Monats-Schätzer: Sie beruhen jeweils auf Daten aus drei aufeinanderfolgenden Befragungsmonaten und rücken stets um einen Monat weiter. Der erste Drei-Monats-Schätzer der berechneten Zeitreihen zu depressiven Symptomen umfasst zum Beispiel Daten aus dem Zeitraum Mitte April bis Mitte Juli 2019. Der nächste Drei-Monats-Schätzer umfasst dann die Monate Mitte Mai bis Mitte August. Alle Drei-Monats-Fenster beginnen und enden mit einer Monatsmitte, weil die monatlichen Erhebungswellen der ausgewerteten Surveys meist etwa zur Monatsmitte begannen. In der zeitlichen Umgebung von Datenlücken werden teilweise Schätzer basierend auf Daten aus zwei Monaten innerhalb eines Drei-Monats-Fensters berechnet, niemals jedoch werden Schätzer basierend auf Daten aus nur einem Monat berechnet.

Vorteil der Verwendung von Drei-Monats-Schätzern: Durch die Berechnung von gleitenden monatlichen Schätzern von Indikatoren der psychischen Gesundheit basierend auf Daten aus jeweils drei (oder zwei) Monaten werden die Stichprobengrößen je Schätzer erhöht und somit auch eine höhere statistische Genauigkeit der Schätzer erreicht. Außerdem trägt dieser Ansatz zur Glättung der Zeitreihen bei, indem zufällige, Stichprobengrößen-bedingte Schwankungen reduziert werden. Bei Ein-Monats-Schätzern käme es zu größeren Fluktuationen durch zufällige Abweichungen zwischen Schätzern, wodurch Trends weniger gut erkennbar bzw. interpretierbar wären. Dies kann besonders bei Indikatoren mit einer geringen Prävalenz und entsprechend geringen Fallzahlen in den Stichproben zum Problem werden sowie bei nach Geschlecht, Alter oder Bildung stratifizierten Zeitreihen.

### **Glättungskurven**

Obwohl die gleitenden Drei-Monats-Schätzer bereits geglättete Zeitreihen bilden, kann es dennoch zu Fluktuationen kommen, die die Interpretation der Zeitreihen erschweren. Damit die allgemeinen Verläufe in den Indikatoren der psychischen Gesundheit besser erkennbar sind, werden daher zusätzlich zu den Drei-Monats-Schätzern Glättungskurven berechnet. Diese beruhen auf „Generalized Additive Models“, die sogenannte „Smoothing Splines“ als Terme enthalten ([Wood, 2003](#)). Sie sollen den zeitlichen Verlauf der Mittelwert- und Anteilsschätzungen kontinuierlich und geglättet approximieren. Die Modellierung basiert hierbei auf wöchentlichen Datenpunkten und resultiert in wöchentlichen Schätzern, durch die eine Kurve gezogen wird. Diese wöchentlichen Schätzer sind im vorliegenden Datensatz enthalten. Nach der aktuellen Methode können die Glättungskurven nur für längere Zeitreihen geschätzt werden. Es können keine Konfidenzintervalle berechnet werden, die die statistische Unsicherheit der Glättungskurven quantifizieren würden. Auch deshalb sollten die Glättungskurven gemeinsam mit den gleitenden Drei-Monats-Schätzern und ihren Konfidenzintervallen betrachtet werden.

## Gewichtung

Die Drei-Monats-Schätzer werden unter Berücksichtigung von Stichprobengewichten mit Surveyprozeduren berechnet (siehe [Allen et al., 2021](#)). Die Stichprobengewichte berücksichtigen in der sogenannten Designgewichtung die unterschiedlichen Ziehungswahrscheinlichkeiten der Teilnehmenden im Rahmen des telefonischen Dual-Frame-Verfahrens (Festnetz, Mobilfunk). Darauf aufbauend wird die Stichprobe aus den RKI-Studien GEDA-EHIS, GEDA und COVIMO an öffentliche Bevölkerungsstatistiken (d.h., aktuelle Bevölkerungsstruktur und Zusammensetzung) nach Alter, Geschlecht, Bildung und Region angepasst. Die Gewichtungsfaktoren werden vom Epidemiologischen Daten- und Befragungszentrum (EDZ) des RKI bereitgestellt.

## Standardisierung

Zusätzlich standardisiert wird in den Berechnungen der Drei-Monats-Schätzer und Glättungskurven nach Alter, Geschlecht und Bildung gemäß CASMIN-Klassifikation ([Brauns et al., 2003](#)). Damit werden Abweichungen in der Stichprobenzusammensetzung zwischen den Erhebungszeitpunkten ausgeglichen. Diese Standardisierung erfolgt unter Einbeziehung des Mikrozensus (Schwarz, 2001) von 2018 als Standardbevölkerung. Als Grundlage für diese Standardisierung werden die Schätzer anhand von Geschlecht, Alter und Bildung modelliert: Drei-Monats-Schätzer und Konfidenzintervalle, die Bevölkerungsmittelwerte von Skalen (z.B. PHQ-2-Mittelwerte) abbilden, beruhen auf linearen Regressionen, die die analysierten Indikatoren der psychischen Gesundheit innerhalb von jedem Drei-Monats-Fenster auf Geschlecht, Alter und Bildung regressieren. Bei dichotomen Merkmalen (z.B. Bevölkerungsanteil mit oder ohne Belastung durch depressive Symptome im auffälligen PHQ-2-Wertebereich) werden äquivalent logistische Regressionen geschätzt. In den linearen Regressionen werden alle Drei-Weg-Interaktionen zwischen Geschlecht, Alter und Bildung berücksichtigt. In den logistischen Regressionen werden dagegen nur Zwei-Weg-Interaktionen berücksichtigt, weil es durch teilweise niedrige Fallzahlen in den Stichproben oder einzelnen Stichproben-Subgruppen bei Drei-Weg-Interaktionen häufiger zu leeren Interaktionszellen käme. Die Berechnung der Glättungskurven über „Generalized Additive Models“ mit „Smoothing Splines“ basiert ebenfalls auf nach Geschlecht, Alter, Bildung und Region gewichteter Modellierung sowie um Geschlecht, Alter und Bildung standardisierten Vorhersagen.

Diese Berechnungsweise ermöglicht bei Stratifizierung der Zeitreihen nach Geschlecht, Alter und Bildung eine weitere Form der Standardisierung: Die jeweiligen Bevölkerungssubgruppen werden einander mit Blick auf die anderen beiden Charakteristika angeglichen. Für Geschlechtervergleiche wird die Alters- und Bildungsverteilung bei Frauen und Männern in der Berechnung angeglichen, für Altersvergleiche die Geschlechter- und Bildungsverteilung und für Bildungsvergleiche die Geschlechter- und Altersverteilung. So kann ausgeschlossen werden, dass sich Unterschiede aus ungleichen Verteilungen der jeweiligen anderen Charakteristika ergeben. Bei der Interpretation standardisierter Schätzer ist zu beachten, dass sie wegen dieser Angleichungen zwischen Gruppen nicht als repräsentative Schätzer für die tatsächlichen Mittelwerte und Anteile der jeweiligen Bevölkerungsgruppen mit ihren realen Zusammensetzungen interpretiert werden können. Der Datensatz enthält neben standardisierten Schätzern daher auch nicht standardisierte Schätzer. Diese lassen sich als

repräsentative Schätzer interpretieren. Allerdings ist bei nicht standardisierten Schätzern nicht auszuschließen, dass sich Unterschiede zwischen Bevölkerungsgruppen aus ungleichen Verteilungen von Alter, Geschlecht oder Bildung ergeben.

## **Fehlende Werte und fehlende Schätzer**

Bei fehlenden Werten in den Indikatoren psychischer Gesundheit werden Beobachtungen fallweise ausgeschlossen. Bei fehlenden Angaben zur Bildung wird das mittlere Bildungsniveau (CASMIN-Klassifikation) zugewiesen. Fehlende Angaben zu Geschlecht oder Alter kommen nicht vor.

Zu Lücken in den Zeitreihen kommt es hauptsächlich wegen Unterbrechungen in der Datenerhebung. Es kann jedoch vor allem in der Schätzung von Bevölkerungsanteilen auch zu fehlenden Drei-Monats-Schätzern kommen, wenn sich wegen zu geringer Fallzahlen in den Interaktionen in den unter „Standardisierung“ beschriebenen Regressionsmodellen leere Zellen ergeben. Wenn zum Beispiel in einem Drei-Monats-Fenster keine 18-29-Jährigen der niedrigen Bildungsgruppe mit Angstsymptomen im auffälligen Wertebereich des GAD-2 vorkommen, dann fehlt der Schätzer für die Anteile mit Angstsymptomen im auffälligen Wertebereich in diesem Zeitraum vollständig und wird im Datensatz mit NA („not available“) ausgewiesen. Da Schätzungen vor und nach Datenlücken häufig auf Daten aus nur zwei Monaten beruhen, kann es wegen geringerer Fallzahlen vor allem in diesen Fällen zu fehlenden Schätzern kommen.

## **Eine ausführliche Beschreibung dieser Methodik ist hier zu finden:**

Junker S, Damerow S, Walther L, Mauz E. Development of a prototype for high-frequency mental health surveillance in Germany: data infrastructure and statistical methods. Front Public Health. 2023; 11:1208515. doi: [10.3389/fpubh.2023.1208515](https://doi.org/10.3389/fpubh.2023.1208515).

## **Aufbau und Inhalt des Datensatzes**

---

Der Datensatz enthält aggregierte Daten aus den Studien GEDA/EHIS 2019-2020, GEDA 2020, COVIMO, GEDA 2021, GEDA 2022 und GEDA 2023, wie sie in der hochfrequenten Mental Health Surveillance aufbereitet wurden. Im Datensatz enthalten sind:

- Daten der hochfrequenten Mental Health Surveillance
- Datensatzdokumentation in deutscher Sprache
- Lizenz-Datei mit der Nutzungslizenz des Datensatzes in Deutsch und Englisch
- Metadaten zur automatisierten Weiterverarbeitung

## **Daten der hochfrequenten Mental Health Surveillance**

Die Daten der hochfrequenten Mental Health Surveillance zu drei Indikatoren der psychischen Gesundheit Erwachsener in Deutschland sind bereitgestellt als:

| [high\\_frequency\\_mental\\_health\\_surveillance.tsv](#)

## Variablen und Variablenausprägungen

Die Daten enthalten folgende Variablen und Variablenausprägungen:

Variablenname	Typ	Ausprägung	Erläuterung
model	string	3M_Mittelwerte , 3M_Anteile , Mittelwerte_smooth , Anteile_smooth	Gibt an, welches Modell zur Schätzung genutzt wurde, eine lineare (3_Monate_Mittelwerte) bzw. logistische Regression (3_Monate_Anteile) für Drei-Monats-Fenster oder ein generalized additive model mit einem smoothing spline unter der Annahme einer Gaussverteilung (Mittelwerte_smooth) oder Binomialverteilung (Anteile_smooth) basierend auf wöchentlichen Datenpunkten
instrument	string	SRMH , PHQ-2 , GAD-2	Gibt das Instrument an, das genutzt wurde und für das der fit (Schätzer) berechnet wurde  SRMH : Self-rated mental health bzw. selbsteingeschätzte psychische Gesundheit  PHQ-2 : (Kurzscreener "Patient Health Questionnaire-2"): depressive Symptome,  GAD-2 (Kurzscreener "Generalized Anxiety Disorder-2"): Angstsymptome



type	string	Summenscore , kategorial	Gibt an, ob das Instrument als Summenscore oder aber kategorial mit einem Cut-Off-Wert genutzt wurde, um darüber und darunter liegende Anteile zu schätzen
category	string	gesamt , Altersgruppe , Bildungsgruppe , Geschlecht	Gibt an, nach welcher Kategorie stratifiziert wurde  Bildungsgruppe : Bildungsgruppe nach CASMIN  Geschlecht : wie bei Geburt zugewiesen
cat_value	string	gesamt , 18-29 Jahre , 30-44 Jahre , 45-64 Jahre , 65+ Jahre , niedrige Bildungsgruppe , mittlere Bildungsgruppe , hohe Bildungsgruppe , männlich , weiblich	Gibt an, für welche Subgruppe der fit (Schätzer) berechnet wurde
standardized	boolean	TRUE , FALSE	Gibt an, ob die Werte bei Stratifizierung für die anderen Kategorien standardisiert wurden
date	date	yyyy-mm-dd	Gibt die Mitte des Drei-Monats-Fensters bzw. der Woche an, für den der fit (Schätzer) berechnet wurde
fit	float	≥0 , NA	Gibt den geschätzten Bevölkerungsmittelwert des Summenscores bzw. den geschätzten Anteil der Bevölkerung an, der über bzw. unter dem Cut-Off-Wert liegt,

			wenn verfügbar bzw. berechenbar, alternativ NA
se	float	>0 , NA	Gibt den Standardfehler an, wenn verfügbar, alternativ NA
lwr	float	≥0 , NA	Gibt das untere 95%-Konfidenzintervall an, wenn verfügbar, alternativ NA
upr	float	≥0 , NA	Gibt das obere 95%-Konfidenzintervall an, wenn verfügbar, alternativ NA
period_duration	string	3_Monate , 2_Monate , 1_Woche	Gibt an, auf welche Periode sich der fit bezieht, auf eine Woche (für Anteil_smooth und Mittelwerte_smooth) oder ob innerhalb des jeweiligen Drei-Monats-Fensters für zwei oder drei Monate Beobachtungen vorhanden waren
period_text	string	Mitte <Monat> bis Mitte <Monat>	Beschreibt explizit das Drei-Monats-Fenster, für das die fits berechnet wurden

## Metadaten

Zur Erhöhung der Auffindbarkeit sind die bereitgestellten Daten mit Metadaten beschrieben. Über GitHub Actions werden Metadaten an die entsprechenden Plattformen verteilt. Für jede Plattform existiert eine spezifische Metadaten-datei, diese sind im Metadatenordner hinterlegt:

[Metadaten/](#)

Versionierung und DOI-Vergabe erfolgt über [Zenodo.org](https://zenodo.org). Die für den Import in Zenodo bereitgestellten Metadaten sind in der [zenodo.json](#) hinterlegt. Die Dokumentation der einzelnen Metadatenvariablen ist unter <https://developers.zenodo.org/representation> nachlesbar.

[Metadaten/zenodo.json](#)

## Hinweise zur Nachnutzung der Daten

---

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

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

## Lizenz

Der Datensatz “Hochfrequente Mental Health Surveillance” ist lizenziert unter der [Creative Commons Namensnennung 4.0 International Public License | CC-BY](#) .

Die im Datensatz bereitgestellten Daten sind, unter Bedingung der Namensnennung des Robert Koch-Instituts als Quelle, frei verfügbar. Das bedeutet, jede Person hat das Recht die Daten zu verarbeiten und zu verändern, Derivate des Datensatzes zu erstellen und sie für kommerzielle und nicht kommerzielle Zwecke zu nutzen. Weitere Informationen zur Lizenz finden sich in der [LICENSE](#) bzw. [LIZENZ](#) Datei des Datensatzes.

## Literatur

---

Ahmad F, Jhaji AK, Stewart DE, Burghardt M, Bierman AS. Single item measures of self-rated mental health: a scoping review. BMC Health Serv Res. 2014; 14:398.

Allen J, Born S, Damerow S, Kuhnert R, Lemcke J, Müller A et al. German Health Update (GEDA 2019/2020-EHIS) – Background and Methodology. J Health Monitor. 2021; 6:66–79.

Brauns H, Scherer S, Steinmann S. The CASMIN Educational Classification in International Comparative Research. In: Hoffmeyer-Zlotnik, J.H.P., Wolf, C. (eds) Advances in Cross-National Comparison. 2003. Springer, Boston, MA.

Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med. 2007; 146:317–25.

Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res. 2005; 58:163–71.

Schwarz N. The German Microcensus. Schmollers Jahrbuch/J Appl Soc Sci. 2001. 121:649-654

Wood SN. Thin Plate Regression Splines. JR Stat Soc, B Stat Methodol. 65(1):95–114.