



# Gesundheitsmonitoring am Robert Koch-Institut – Auswirkungen eines veränderten Studiendesigns auf Stichprobenzusammensetzung und Prävalenzschätzer durch den Start des Panels „Gesundheit in Deutschland“

**Autorinnen und Autoren:** Elvira Mauz\*, Felicitas Vogelgesang\*, Stefan Damerow, Anja Schienkiewitz, Niels Michalski, Beate Gaertner, Jennifer Allen, Jens Baumert, Yong Du, Ronny Kuhnert, Johannes Lemcke, Ramona Scheufele, Angelika Schaffrath Rosario

**Institution:** Robert Koch-Institut, Abteilung für Epidemiologie und Gesundheitsmonitoring, Berlin

\*geteilte Erstautorenschaft

## Abstract

**Hintergrund:** Das Robert Koch-Institut (RKI) beobachtet mit wiederholten Datenerhebungen kontinuierlich zentrale Gesundheitsindikatoren in der Allgemeinbevölkerung. Da Änderungen im Erhebungsdesign die Prävalenzschätzungen beeinflussen und somit eine Trendinterpretation erschweren können, wurde der Start des RKI-Panels „Gesundheit in Deutschland“ durch eine Methodenstudie begleitet.

**Methode:** Das RKI-Panel beruht auf einer Zufallsstichprobe aus Einwohnermeldeamtsdaten und selbstadministrierter schriftlicher Befragung (online oder Papier). Stichprobenzusammensetzung, Prävalenzschätzer und Antwortverhalten wurden mit parallel erhobenen Daten des Telefonsurveys GEDA 2024 verglichen. Zur Quantifizierung methodenbedingter Unterschiede in den Prävalenzschätzern wurden Daten früherer Surveys zur Modellierung einbezogen.

**Ergebnisse:** Im RKI-Panel 2024 konnten junge Erwachsene, Hochaltrige und Personen mit niedriger Bildung besser erreicht werden. Die Prävalenzschätzungen weichen teils deutlich von GEDA 2024 ab, insbesondere bei Indikatoren der psychischen Gesundheit und deren Einflussfaktoren. Im RKI-Panel zeigt sich bei Hochaltrigen häufiger eine eingeschränkte körperliche, bei jungen Erwachsenen eine schlechtere psychische Gesundheit. Über die Modellierung kann meist – trotz methodenbedingter Prävalenzunterschiede – der Trendverlauf beurteilt werden.

**Schlussfolgerungen:** Das RKI-Panel bildet die Bevölkerung in Deutschland realistischer ab als die bisherigen Telefonsurveys. Prävalenzabweichungen begründen sich in Effekten des Erhebungsmodus, des Fragebogendesigns sowie veränderter Stichprobenzusammensetzung.

**Keywords:** Surveydesign, Surveillance, Methodenwechsel, Prävalenzvergleich, Antwortverhalten, Repräsentativität, Verzerrung, Korrektur

## Informationen zu Artikel und Zeitschrift

Eingereicht: 14.08.2025

Akzeptiert: 29.10.2025

Veröffentlicht: 05.12.2025


Artikel peer reviewed

Zitierweise: Mauz E, Vogelgesang F, Damerow S, Schienkiewitz A, Michalski N, Gaertner B, et al. Gesundheitsmonitoring am Robert Koch-Institut – Auswirkungen eines veränderten Studiendesigns auf Stichprobenzusammensetzung und Prävalenzschätzer durch den Start des Panels „Gesundheit in Deutschland“. J Health Monit. 2025;10(4):e 13558. doi: 10.25646/13558

Angelika Schaffrath Rosario  
[rosarioa@rki.de](mailto:rosarioa@rki.de)

Robert Koch-Institut, Berlin  
Journal of Health Monitoring  
[www.rki.de/jhealthmonit](http://www.rki.de/jhealthmonit)

Englische Version des Artikels  
[www.rki.de/jhealthmonit-en](http://www.rki.de/jhealthmonit-en)

 Open access



[CC BY 4.0 Lizenzvertrag](https://creativecommons.org/licenses/by/4.0/)  
[Namensnennung 4.0 International](https://creativecommons.org/licenses/by/4.0/)

Gesundheitsberichterstattung des Bundes.  
Gemeinsam getragen von RKI und Destatis.



Das Robert Koch-Institut ist ein  
Bundesinstitut im Geschäftsbereich des  
Bundesministeriums für Gesundheit

## 1. Einleitung

Das Robert Koch-Institut (RKI) erhebt im Rahmen seines Gesundheitsmonitorings kontinuierlich zentrale Indikatoren zur Beschreibung der gesundheitlichen Lage der Allgemeinbevölkerung [1]. Dazu gehören sowohl einmalige Erhebungen zu bislang unzureichend erforschten Gesundheitsthemen als auch der systematische Aufbau von Zeitreihen ausgewählter Indikatoren, die eine Beurteilung ihrer Entwicklung in verschiedenen Bevölkerungsgruppen ermöglichen. Seit 2024 werden regelmäßig aktualisierte Ergebnisse auf dem Webportal der Gesundheitsberichterstattung (GBE) veröffentlicht und sukzessive um weitere Indikatoren ergänzt [2].

Um sowohl kontinuierlich als auch in Krisensituationen zuverlässige Daten für die Gesundheitsberichterstattung bereitzustellen, wurde im Jahr 2024 am RKI eine neue Panel-Infrastruktur eingeführt [3]. Mit dem RKI-Panel „Gesundheit in Deutschland“ werden zentrale Herausforderungen epidemiologischer Erhebungen wie sinkende Teilnahmequoten, steigende Erhebungskosten und die erschwerte Erreichbarkeit bestimmter Bevölkerungsgruppen, insbesondere Hochaltriger und Personen mit niedriger Bildung, adressiert [4–9]. Ziel ist es, regelmäßige Erhebungen von Befragungs-, Untersuchungs- und Labordaten langfristig mit konstanter Methodik zu gewährleisten und in Krisensituationen schnell und flexibel erforderliche Befragungen umsetzen zu können [3]. Vor Einführung des RKI-Panels wurden seit 1984 verschiedene Gesundheitssurveys durchgeführt [1], die sich im Zeitverlauf sowohl im Stichproben- als auch im Erhebungsdesign unterschieden. Während einige Surveys auf Zufallsstichproben aus Einwohnermeldeämtern in bundesweit verteilten Erhebungsstellen (EMA-Stichproben) basierten, nutzten andere Telefonstichproben aus Festnetz- bzw. später aus Festnetz- und Mobilfunknummern (Dual-Frame-Stichproben). Die Datenerhebung erfolgte teils telefonisch, teils schriftlich (online oder Papier) (Tabelle 1). Das RKI-Panel basiert, wie zuletzt die Studie GEDA 2014/2015-EHIS [10], auf einer EMA-Stichprobe und einem schriftlichen Mixed-Mode-Design (Online- bzw. Papierbefragung) [3].

Für die valide Interpretation zeitlicher Trends ist eine möglichst konstante Erhebungsmethodik essenziell, da methodische Änderungen die Vergleichbarkeit von Prävalenzschätzungen beeinträchtigen und zu abweichenden Ergebnissen führen können [11–13]. Diese Methodeneffekte können sich sowohl auf die Teilnahmebereitschaft und damit auf die Repräsentativität der Stichprobe (Selektionseffekte) als auch auf das Antwortverhalten und somit auf die Schätzung der Häufigkeit eines Merkmals auswirken (Mess-effekte; siehe Überblick in [14, 15]). Die Umstellung des Studiendesigns von den GEDA-Studien („Gesundheit in Deutschland Aktuell“) der letzten Jahre auf das RKI-Panel beinhaltet mehrere Änderungen: (1) die Umstellung von einer Dual-Frame-Telefonstichprobe auf eine EMA-Stichprobe, (2) den Wechsel des Erhebungsmodus von telefonischen Interviews zu einer selbstadministrierten schriftlichen Mixed-Mode-Erhebung, (3) ein aktualisiertes Gewichtungungsverfahren sowie (4) die thematische Aufteilung des Fragebogens auf vier Module. Entsprechend wurde von Effekten auf die Stichprobenstruktur, das Antwortverhalten und damit auf die Prävalenzschätzungen ausgegangen.

Um trotz des Methodenwechsels die begonnenen Zeitreihen im Webportal der GBE fortzuführen und verlässlich interpretieren zu können, wurden begleitend zum Start des RKI-Panels methodische Untersuchungen durchgeführt: (1) Parallel zur ersten Jahreserhebung des RKI-Panels (RKI-Pa-

### Kernaussagen

- ▶ Die Methodeneffekte des Wechsels von der GEDA-Studie zum RKI-Panel konnten durch eine Methodenstudie beurteilt werden.
- ▶ Der Methodenwechsel kann insbesondere bei Indikatoren der psychischen Gesundheit die Prävalenzen deutlich verändern.
- ▶ Mit dem RKI-Panel kann die Bevölkerungsstruktur in Deutschland besser abgebildet werden als mit den bisherigen GEDA-Studien.
- ▶ Bevölkerungsgruppen wie junge Erwachsene, Hochaltrige ab 80 Jahren oder Personen mit niedriger Bildung sind im RKI-Panel besser repräsentiert.
- ▶ Themen mit Tendenz zu sozial erwünschten Antworten, subjektive Einschätzungen sowie graduelle Antwortmöglichkeiten (Likert-Skalen) sind besonders anfällig für Methodeneffekte.

nel 2024) fand im Sommer 2024 eine telefonische GEDA-Studie (GEDA 2024) statt, die den Vergleich von Stichprobenzusammensetzung und Prävalenzschätzungen identisch erhobener Indikatoren zwischen dem „alten“ und dem „neuen“ Studiendesign ermöglicht. (2) Ergänzend wurden Daten früherer Surveys herangezogen, um die Auswirkungen des Designwechsels auf Gesundheitsindikatoren auch im Zeitverlauf zu identifizieren, zu beschreiben und, soweit möglich, zu quantifizieren.

Ziel der vorliegenden Arbeit ist es, die methodischen Effekte des Wechsels im Erhebungsdesign einzuschätzen und zu quantifizieren, um eine verlässliche Interpretation der Zeitreihen der Gesundheitsindikatoren zu gewährleisten. Durch die Entwicklung von Korrekturfaktoren sollen Prävalenzschätzungen aus unterschiedlichen Surveydesigns vergleichbar werden. Für jeden im Webportal der Gesundheitsberichterstattung dargestellten Indikator wird ein Indikatorenblatt bereitgestellt, welches die Auswirkungen der neuen Erhebungsmethodik dokumentiert und entsprechend angepasste Prävalenzen ausweist.

## 2. Methode

Die für die methodischen Analysen verwendeten RKI-Studien wurden anhand der wesentlichen Unterschiede in der Stichprobenziehung und dem Erhebungsmodus in zwei Studientypen eingeteilt (Tabelle 1):

1. Studien mit einer EMA-Stichprobe und einem schriftlichen Mixed-Mode-Design wurden dem Studientyp 1 zugeordnet. Dazu zählt das neu etablierte RKI-Panel.
2. Studien basierend auf einer Telefonstichprobe (Festnetz- oder Dual-Frame-Stichprobe) und Telefoninterviews wurden dem Studientyp 2 zugeordnet.

## 2.1 Datengrundlagen

### Panel „Gesundheit in Deutschland“ – Rekrutierungsstudie und RKI-Panel 2024 (Studientyp 1)

Im Rahmen einer Rekrutierungsstudie wurde im Jahr 2024 das RKI-Panel „Gesundheit in Deutschland“ etabliert. Die Grundgesamtheit für die Rekrutierung waren die in der Bundesrepublik Deutschland im Erhebungszeitraum lebenden Personen ab 16 Jahren, unabhängig von Staatsangehörigkeit oder Geburtsland. Allerdings wurden nur deutschsprachige Erhebungsinstrumente eingesetzt. Die Stichprobenziehung basierte auf einer zweifach geschichteten Zufallsauswahl: Aus der Gesamtheit aller Gemeinden in Deutschland wurden unter Beachtung der regionalen Struktur zufällig 359 Primary Sampling Units, sogenannte Sample Points, gezogen (erste Auswahlstufe). In der zweiten Auswahlstufe wurden für jeden Sample Point, stratifiziert nach Altersgruppen, Adressen aus den Adressregistern der jeweiligen Einwohnermeldeämter gezogen. Die rund 167.000 ausgewählten Personen wurden zur Teilnahme an einer kurzen Befragung eingeladen und um Einwilligung zur Teilnahme an künftigen Befragungen im Rahmen des Panels gebeten [3]. Die Rekrutierungsquote lag bei 29 %. Die Befragung erfolgte bei Personen unter 70 Jahren in einem schriftlichen sequenziellen Mixed-Mode-Design, bei dem mit dem Einladungsschreiben ein Link und QR-Code zur Onlinebefragung zugesandt wurden. Bei Nichtteilnahme wurde im Zuge des zweiten Erinnerungsschreibens neben der Onlineteilnahme auch ein Papierfragebogen angeboten. Bei Personen ab 70 Jahren wurde auf ein simultanes Mixed-Mode-Design umgestellt, bei dem bereits mit der Einladung zur Studie zwischen Online- und Papierfragebogen gewählt werden konnte. Dieses Vorgehen wurde auch in den bis zu zwei Erinnerungsschreiben fortgesetzt [16].

Für die erste Jahresherhebung 2024 umfasste das RKI-Panel 46.863 registrierte Teilnehmende ab einem Alter von 18 Jahren: 24.881 Frauen, 21.856 Männer, 126 Personen mit anderer Geschlechtsidentität. Im Jahr 2024 wurden diese im Abstand von ca. zwei Monaten zu drei Zeitpunkten (Teilwellen; Quartal 2, 3 und 4) zu Gesundheitsbefragungen eingeladen. Insgesamt gab es vier verschiedene Fragebögen (A, B, C, D) zu unterschiedlichen Themenbereichen. Die vier Fragebögen sind thematisch gruppiert und die Themen damit weniger durchmischt als in bisherigen RKI-Surveys. Die im Panel registrierten Personen erhielten in einem vorher definierten rotierenden Verfahren in jeder der drei Teilwellen

einen der vier Fragebögen, sodass jede Person drei der vier Fragebögen erhielt. Zusätzlich enthielten alle in der ersten Teilwelle am Ende des Fragebogens vertiefte Fragen zu soziodemografischen Informationen. Die Teilnahme war online und schriftlich-postalisch möglich, wobei die in der Rekrutierungsbefragung realisierte Methode die Erhebungsmethode für die Jahreswelle 2024 bestimmte. Die Datenerhebung begann im Mai 2024 mit der ersten Teilwelle und wurde Anfang Januar 2025 mit der dritten Teilwelle abgeschlossen. Die Wiederteilnahmequote (Anteil der Teilnehmenden im Verhältnis zu den im Panel registrierten Personen) lag in den einzelnen Teilwellen gemäß den Standards der American Association for Public Opinion Research (AAPOR [17]) zwischen 75 % und 81 %. Eine ausführliche Darstellung der Methodik und Responsequoten (auch stratifiziert nach Alter und Geschlecht) findet sich an anderer Stelle [18].

### Datengrundlagen der Methodenstudie: Parallelerhebung GEDA 2024 (Studientyp 2) und Sommermonate (Q2 + Q3) des RKI-Panels 2024 (Studientyp 1)

Die ersten beiden Teilwellen des RKI-Panels 2024 (Q2 und Q3) wurden von der Methodenstudie GEDA 2024 im gleichen Design der bisherigen GEDA-Telefonsurveys begleitet. GEDA 2024 beruht auf einer Telefonstichprobe von Personen ab 18 Jahren und erfolgte mittels Dual-Frame-Ansatz über Telefoninterviews. Zum Einsatz kam ein gegenüber den vorangegangenen GEDA-Wellen deutlich reduziertes Befragungsinventar, wobei die ausgewählten Inhalte identisch zum RKI-Panel 2024 erhoben wurden (siehe 2.2). Die Erhebung erfolgte zwischen Juni und September 2024 mit jeweils etwa 1.000 Teilnehmenden pro Monat ( $n=4.016$ , Response-Quote 20 %). Für die Vergleichsanalysen wurden die Daten des RKI-Panels 2024 aus den parallel erhobenen ersten zwei von drei Teilwellen herangezogen (Quartal 2 und 3,  $n \approx 18.000$  in Abhängigkeit vom jeweiligen Fragebogen).

### Trendanalysen: Vorangegangene Surveys des RKI-Gesundheitsmonitorings

Für die Beurteilung der Methodeneffekte über den Zeitverlauf und die modellbasierte Berechnung von Korrekturfaktoren wurden zusätzlich weitere RKI-Befragungssurveys einbezogen und den beiden zu vergleichenden Studientypen zugeordnet:

Studientyp 1: Diesem Studientyp wurde GEDA 2014/2015-EHIS zugeordnet mit einer über die Einwohnermeldeämter erfolgten Stichprobenziehung und einem bereits im Vorfeld umfassend getesteten sequenziellen schriftlichen Mixed-Mode-Design [19, 20], bei dem zunächst ausschließlich ein Onlinefragebogen angeboten wurde und erst im Rahmen einer Erinnerung zusätzlich ein Papierfragebogen. Damit ähnelt das Design von GEDA 2014/2015-EHIS dem des RKI-Pa-

nels, lediglich mit dem Unterschied, dass im RKI-Panel bei Personen ab 70 Jahren direkt ein simultanes Mixed-Mode-Design angeboten wurde.

Studientyp 2: Alle weiteren Befragungsstudien des Gesundheitsmonitorings entsprechen dem Studientyp 2 mit Telefonstichproben und telefonischen Interviews. Die Surveys bis 2012 (GSTel 2003 und 2006; GEDA 2009, 2010, 2012) beruhten ausschließlich auf Festnetzstichproben. Ab 2019/2020 kamen Dual-Frame-Stichproben, d. h. Festnetz und Mobiltelefon, zum Einsatz (2019/2020, 2022 und 2023).

### Gewichtung in den RKI-Surveys

Für jeden RKI-Survey wurde ein Gewichtungsfaktor für die Auswertung erstellt, der das jeweilige Stichprobendesign berücksichtigt sowie selektive Teilnahme so weit wie möglich ausgleicht.

Gewichtung RKI-Panel: Um Verzerrungen aufgrund selektiver Teilnahme sowie Abweichungen der Stichprobe von der Bevölkerungsstruktur so weit wie möglich zu korrigieren, wurde ein mehrstufiges Stichprobengewicht berechnet. Es berücksichtigt zunächst das Stichprobengewicht der initialen Rekrutierungsstudie. Zusätzlich wurden in einer Drop-out-Gewichtung mit den Angaben aus der Rekrutierungsstudie Ausfallgewichte berechnet, um eine selektive Teilnahme an den wiederholten Teilwellen auszugleichen. Abschließend wurden Anpassungen an die Bevölkerungsfortschreibung zum Stand 31.12.2023 bzw. an den Mikrozensus 2021 berechnet. Dabei wurden Alter, Geschlecht, Bundesland, BIK-Gemeindetyp [21], Bildungsgruppen (nach Comparative Analysis of Social Mobility in Industrial Nations, CASMIN [22]) und Haushaltsgröße (Ein- vs. Mehrpersonenhaushalt) berücksichtigt. Die Angaben zur Bildung wurden im Rahmen der Rekrutierungsstudie erhoben. Die Angaben zur Haushaltsgröße stammen aus dem Soziodemografie-Fragebogen in der ersten Teilwelle der Jahreserhebung; falls sie dort fehlten (etwa wegen Nicht-Teilnahme an der ersten Teilwelle), wurden sie durch die entsprechende Angabe aus der Rekrutierungsstudie ersetzt. Die Gewichtung wurde für jede Fragebogenvariante separat berechnet. Eine ausführliche methodische Darstellung erfolgt in einem gesonderten Fachartikel [23].

Gewichtung vorangegangene Surveys: Alle vorangegangenen RKI-Surveys wurden bereits für frühere Auswertungen gewichtet (Details siehe [Annex Tabelle 1](#)). Analog dazu wurde die Parallelerhebung GEDA 2024 nach Alter, Geschlecht, Bundesland, Kreistyp und Bildung gewichtet (zum Bevölkerungsstand 31.12.2023 und gemäß Mikrozensus 2021). Die bisherige Gewichtung unterschied sich gegenüber der Gewichtung im RKI-Panel dahingehend, dass (1) die Gewichtung nach Haushaltsgröße erst im RKI-Panel neu eingeführt wurde und (2) die Bildungsverteilung im RKI-Panel nach der CASMIN-Klassifikation [22] angepasst wurde [24],

während in vorherigen RKI-Surveys nach der ISCED-Klassifikation (International Standard Classification of Education [25]) angepasst wurde.

### 2.2 Gesundheitsindikatoren

Insgesamt wurden elf Indikatoren auf Methodeneffekte untersucht. Diese Indikatoren wurden sowohl im RKI-Panel 2024 als auch in früheren RKI-Surveys als Selbstangaben erhoben und ermöglichen somit Tendaussagen ([Tabelle 1](#)). Für sieben dieser Indikatoren liegen zusätzlich Daten aus der Parallelerhebung GEDA 2024 vor. Die Indikatoren waren im RKI-Panel 2024 in zwei verschiedenen Fragebögen enthalten: Fragebogen A fokussierte auf Indikatoren zur körperlichen Gesundheit und Risikofaktoren, Fragebogen C auf Indikatoren der psychischen Gesundheit und deren Einflussfaktoren sowie Müdigkeit und Erschöpfung.

Um Hypothesen zur erwarteten Richtung der Prävalenzunterschiede zwischen interviewbasierten und selbst-administrierten Erhebungen aufzustellen, wurden die Indikatoren auf Basis aus der Literatur bekannter Einflussfaktoren auf das Antwortverhalten (z. B. [19, 26–28]) nach (1) Grad sozialer Erwünschtheit, (2) Fragetypus (faktisch vs. subjektiv) und (3) Skalenniveau des Antwortformats klassifiziert ([Annex Tabelle 2](#)).

#### Indikatoren zur körperlichen Gesundheit und Risikofaktoren Allgemeiner Gesundheitszustand

Mit dem Minimum European Health Module (MEHM) [29, 30] als festem Bestandteil des RKI-Gesundheitsmonitorings wurden die folgenden drei Indikatoren erfasst:

Der Indikator *Selbsteingeschätzte allgemeine Gesundheit* wurde definiert als der Anteil der Erwachsenen mit einem sehr guten oder guten selbst eingeschätzten allgemeinen Gesundheitszustand. Dies wurde mit der Frage „Wie ist Ihr Gesundheitszustand im Allgemeinen?“ (Antwortmöglichkeiten: Sehr gut | Gut | Mittelmäßig | Schlecht | Sehr schlecht) erfasst.

Der Indikator *Chronisches Kranksein* wurde mit der Frage erfasst: „Haben Sie eine chronische Krankheit oder ein lang andauerndes gesundheitliches Problem? Damit gemeint sind Krankheiten oder gesundheitliche Probleme, die mindestens 6 Monate andauern oder voraussichtlich andauern werden“ (Antwortmöglichkeiten: Ja | Nein).

Der Indikator *Einschränkungen bei Alltagsaktivitäten* wurde definiert als der Anteil der Erwachsenen, die durch ein gesundheitliches Problem seit mindestens sechs Monaten mäßig oder stark bei Tätigkeiten des normalen Alltagslebens eingeschränkt sind. Die Frage lautete: „Sind Sie durch ein gesundheitliches Problem bei Tätigkeiten des normalen Alltagslebens eingeschränkt?“ (Antwortmöglichkeiten: Stark eingeschränkt | Mäßig eingeschränkt | Nicht eingeschränkt).

Tabelle 1: Übersicht über die Datengrundlagen der analysierten Indikatoren

Survey	Stichprobe		Erhebungs- modus	Indikator										
	Einwohnermeldeamts( EMA)-Stichprobe	Telefonstichprobe: Festnetz (F)/ Dual-Frame (DF)	Mixed-Mode schriftlich (webbasiert 🖥️/Papier 📝) /Telefoninterview ①	Selbsteingeschätzte allgemeine Gesundheit	Chronisches Kranksein	Einschränkungen bei Alltagsaktivitäten	Körperliche Einschränkungen	Diabetes mellitus	Adipositas	Rauchen	Selbsteingeschätzte psychische Gesundheit	Depressive Symptomatik	Angstsymptome	Soziale Unterstützung
GSTel03 (2003)		F	①	X				X	X	X				
GSTel06 (2006)		F	①	X				X	X	X				
GEDA 2009		F	①	X					X	X				X
GEDA 2010		F	①	X				X	X	X				X
GEDA 2012		F	①	X				X	X	X				X
GEDA 2014/2015-EHIS	EMA		🖥️📝	X	X		X		X	X		X		X
GEDA 2019/2020-EHIS		DF	①	X	X	X	X		X	X		X		X
GEDA 2022		DF	①	X	X	X		X	X	X	X	X	X	X
GEDA 2023		DF	①	X	X	X			X	X	X	X	X	X
RKI-Panel 2024 Rekrutierungsstudie	EMA		🖥️📝	X	X				X	X	X			
GEDA 2024		DF	①	X	X				X	X	X	X	X	
RKI-Panel 2024	EMA		🖥️📝	X	X	X	X	X	X	(X)*	X	X	X	X

Anmerkungen: EMA=Einwohnermeldeamt, DF=Dual Frame=Auswahl beruhend auf Festnetz und Mobiltelefon, GSTel=telefonische Gesundheitssurveys, GEDA=Gesundheit in Deutschland aktuell, EHIS=European Health Interview Survey.  
\*Die Frage zum Rauchen wurde im RKI-Panel 2024 nur Personen ohne Diabetes gestellt, um einen Diabetesrisikoscore zu ermitteln. Bei Personen mit Diabetes wurde daher die Angabe zum Rauchstatus aus der Rekrutierungsstudie übernommen.

Eingeschränkte Personen wurden zusätzlich gefragt, wie lange die Einschränkungen bereits andauern (Antwortmöglichkeiten: Weniger als 6 Monate | 6 Monate und länger).

Körperliche Einschränkungen

Der Indikator *Körperliche Einschränkungen* bildet schwerwiegende Mobilitätseinschränkungen ab und wurde definiert als der Anteil der Erwachsenen mit großen Schwierigkeiten beim Laufen oder Treppensteigen oder der Unfähigkeit dies zu tun. Dies wurde mit zwei Fragen aus dem European Health Interview Survey [31] erhoben: „Haben Sie Schwierigkeiten einen halben Kilometer (500 Meter) auf ebenem Gelände ohne Gehhilfe zu gehen?“ und „Haben Sie Schwierigkeiten eine Treppe mit 12 Stufen hinauf- oder hinabzusteigen?“ (Antwortmöglichkeiten: Keine Schwierigkeiten | Einige Schwierigkeiten | Große Schwierigkeiten | Es ist mir nicht möglich/Ich bin dazu nicht in der Lage). Eine Mobilitätseinschränkung besteht, wenn bei mindestens einer der beiden Mobilitätsformen große Schwierigkeiten bestehen oder sie gar nicht möglich ist.

Diabetes mellitus

Der Indikator *Diabetes mellitus* ist definiert als der Anteil der Erwachsenen mit einem jemals ärztlich diagnostizierten Diabetes. Ausgenommen sind Frauen, die ausschließlich die Diagnose eines Schwangerschaftsdiabetes berichtet haben. Zur Erfassung des Diabetes mellitus wurde den Teilnehmenden in den unterschiedlichen Surveys in fast identischer Formulierung die Frage gestellt, ob bei ihnen jemals von einer Ärztin/einem Arzt eine Zuckerkrankheit oder Diabetes festgestellt wurde (Antwortmöglichkeiten: Ja | Nein).

Adipositas

Eine *Adipositas* liegt bei einem Body-Mass-Index (BMI)  $\geq 30\text{kg/m}^2$  vor [32]. Der BMI wird berechnet aus Selbstangaben der Teilnehmenden zu Körpergewicht und Körpergröße. Über die einzelnen Erhebungszeitpunkte hinweg variiert die Frageformulierung leicht, die Antworten werden für die Körpergröße in cm und für das Körpergewicht in kg erfasst [33].



## Rauchen

Der Indikator *Rauchen* basiert auf der Frage nach dem aktuellen Rauchstatus von Tabakprodukten mit leicht abweichender Frageformulierung über die Surveys hinweg aber stets gleichbleibenden Antwortmöglichkeiten [33]: Ja, täglich | Ja, gelegentlich | Nein, nicht mehr | Habe noch nie geraucht. Üblicherweise wird dabei die Verbreitung des aktuellen Rauchens (Ja, täglich | Ja, gelegentlich) abgebildet. Im RKI-Panel 2024 wurde diese Frage nur Personen ohne Diabetes mellitus gestellt. Für Personen mit Diabetes wurde die Angabe aus der Rekrutierungsstudie genutzt.

## Indikatoren zur psychischen Gesundheit und deren Einflussfaktoren

### Selbsteingeschätzte psychische Gesundheit

Der Indikator *Selbsteingeschätzte psychische Gesundheit* ist definiert als der Anteil der Erwachsenen, die ihre eigene psychische Gesundheit im Allgemeinen als ausgezeichnet oder sehr gut einschätzen [34]. Eingesetzt wird ein international etabliertes Einzelitem: „Wie würden Sie Ihre psychische Gesundheit im Allgemeinen beschreiben?“ (Antwortmöglichkeiten: Ausgezeichnet | Sehr gut | Gut | Weniger gut | Schlecht).

### Depressive Symptomatik

Depressive Symptome in der Bevölkerung werden mit dem Patient Health Questionnaire-8 (PHQ-8) erfasst [35]. Die Teilnehmenden werden gefragt, wie oft/häufig sie sich jeweils in den letzten zwei Wochen durch acht erfragte Symptome einer depressiven Störung beeinträchtigt gefühlt haben. Den Antwortmöglichkeiten der acht Items sind jeweils Werte zugeordnet: Überhaupt nicht (Wert 0) | An einzelnen Tagen (Wert 1) | An mehr als der Hälfte der Tage (Wert 2) | Beinahe jeden Tag (Wert 3). Diese werden zu einem Skalensummenwert von 0 bis 24 addiert. Das Vorliegen einer *Depressiven Symptomatik* wird ab einem Skalensummenwert von mindestens 10 angenommen.

### Angstsymptome

Der Indikator *Angstsymptome* in der Bevölkerung wird mit dem Inventar Generalized Anxiety Disorder-2 (GAD-2) erfasst [36]. Die Teilnehmenden werden gefragt, wie oft/häufig sie sich in den letzten zwei Wochen durch zwei Kernsymptome einer generalisierten Angststörung beeinträchtigt gefühlt haben. Die Antwortmöglichkeiten sind analog zum PHQ-8 und addieren sich zu einem Skalensummenwert von 0 bis 6. Das Vorliegen einer auffälligen Belastung durch Angstsymptome wird ab einem Skalensummenwert von 3 angenommen.

### Soziale Unterstützung

Der Indikator *Soziale Unterstützung* wird mittels der Oslo-3-Items-Social-Support-Scale (OSSS-3, Oslo-3-Skala) erfasst,

die die subjektiv wahrgenommene Verfügbarkeit sozialer Unterstützung misst [37]. Konkret wurden die Fragen gestellt, „Wie viele Menschen stehen Ihnen so nahe, dass Sie sich auf sie verlassen können, wenn Sie ernste persönliche Probleme haben?“ (Antwortmöglichkeiten: Keine | 1 bis 2 | 3 bis 5 | 6 oder mehr), „Wie viel Anteilnahme und Interesse zeigen andere Menschen an dem, was Sie tun?“ (Antwortmöglichkeiten: Sehr viel | Viel | Weder viel noch wenig | Wenig | Keine) und „Wie einfach ist es für Sie, praktische Hilfe von Nachbarn zu erhalten, wenn Sie diese benötigen?“ (Antwortmöglichkeiten: Sehr einfach | Einfach | Möglich | Schwierig | Sehr schwierig). Durch Addition der Einzelpunktwerte aus den drei Fragen wird ein Index gebildet, der Werte zwischen 3 und 14 Punkten annehmen kann. Der Bereich von 12 bis 14 Punkten ist als starke Unterstützung klassifiziert.

## 2.3 Statistische Analyse der Parallelerhebungen 2024

Die beiden gleichzeitig durchgeführten Erhebungen GEDA 2024 und RKI-Panel 2024 (eingeschränkt auf Quartal 2 und 3) wurden hinsichtlich (1) der Stichprobenzusammensetzung, (2) der Prävalenz der Gesundheitsindikatoren und (3) des Antwortverhaltens verglichen. Es gingen nur Personen ab 18 Jahren in die Auswertung ein. Die Auswertungen erfolgten gewichtet und für das RKI-Panel unter Berücksichtigung der Clusterung in den Sample Points mit den Survey-Prozeduren von SAS 9.4 (SAS Institute, Cary, NC, USA). Für das RKI-Panel 2024 wurden die Gewichtungsfaktoren genutzt, die für die gesamte Jahreswelle erstellt wurden (ohne Einschränkung auf Quartal 2 und 3). Beide Erhebungen sind auf den Bevölkerungsstand 31.12.2023 gewichtet. Für stratifizierte Auswertungen nach Bildung wurde innerhalb der Bildungsgruppen eine Standardisierung anhand der europäischen Standardbevölkerung [38] vorgenommen, so dass in diesen Auswertungen allen Bildungsgruppen dieselbe Alters- und Geschlechtsverteilung zu Grunde liegt.

### Stichprobenzusammensetzung

Zur Beurteilung der Stichprobenzusammensetzung wurden gewichtete und ungewichtete Anteile sowie der Anteil fehlender Werte für Geschlecht (basierend auf aktueller Geschlechtsidentität), fünf Altersgruppen (18–29, 30–44, 45–64, 65–79, ab 80 Jahren), drei Bildungsgruppen (nach CASMIN-Klassifikation: niedrig, mittel, hoch [22]) sowie Haushaltsgröße in beiden Erhebungen gegenübergestellt. Um zu testen, ob sich mögliche Differenzen in der Stichprobenzusammensetzung durch die veränderte Gewichtung im RKI-Panel erklären lassen, wurden in einer Sensitivitätsanalyse die GEDA-2024-Daten wie das RKI-Panel gewichtet (d. h. mit Berücksichtigung der Haushaltsgröße und auf Basis der CASMIN-Bildungsklassifikation anstelle der ISCED-Bildungsklassifikation).

### Prävalenzvergleich der Gesundheitsindikatoren

Für den Vergleich der Prävalenzschätzer werden die gewichteten Anteile für Gesamt sowie stratifiziert nach Geschlecht, Alters- sowie Bildungsgruppen (nach CASMIN-Klassifikation) dargestellt. Beim Vergleich konnte nicht unterschieden werden zwischen Effekten aufgrund einer veränderten Stichprobenzusammensetzung, aufgrund eines veränderten Erhebungsmodus oder einer veränderten Fragebogengestaltung. Effekte aufgrund der veränderten Gewichtung wurden in einer Sensitivitätsanalyse geprüft, in der GEDA 2024 wie das RKI-Panel gewichtet wurde.

### Antwortverhalten in Abhängigkeit vom Erhebungsmodus und Antwortformat

Die erhobenen Indikatoren unterscheiden sich hinsichtlich ihres Frage- und Antwortformats, die wiederum unterschiedlich stark durch den Wechsel des Erhebungsmodus beeinflusst werden [26, 27, 39]. Daher wurden die Indikatoren anhand des Formats der ihnen zugrunde liegenden Einzelitems klassifiziert (kategoriales, einfaches oder mehrfaches Likert-Skalen-Format oder metrisches Antwortformat) und einander gegenübergestellt. Darüber hinaus wurden die Anteile fehlender Werte in Abhängigkeit vom Antwortformat verglichen.

### 2.4 Analyse von zeitlichen Trends in der Prävalenzentwicklung und Schätzung des Methodeneffekts

Unterschiede, die sich im direkten Prävalenzvergleich der Parallelerhebungen GEDA 2024 und RKI-Panel 2024 zeigen, sind aufgrund der geringeren Fallzahl von GEDA 2024 gerade in Subgruppen mit größerer statistischer Unsicherheit behaftet. Daher wurden Prävalenzschätzer aus vorangegangenen Surveys in die Analyse mit einbezogen (Tabelle 1, Annex Tabelle 1), um die Größe des Methodeneffekts abzuschätzen und so den Trendverlauf besser beurteilen zu können.

Bei der Schätzung des Methodeneffekts konnte nicht zwischen Effekten aufgrund einer veränderten Stichprobenzusammensetzung, aufgrund eines veränderten Erhebungsmodus, einer veränderten Fragebogengestaltung oder aufgrund veränderter Gewichtung unterschieden werden. Stattdessen wurde der Gesamt-Methodeneffekt geschätzt, indem die Prävalenztrends zwischen den beiden Studientypen verglichen wurden (d. h. Studientyp 1 vs. Studientyp 2). Es wurde für jeden Indikator ein Modell mittels logistischer Regression aufgestellt, in das als unabhängige Variablen der Erhebungszeitpunkt (mittlerer Befragungszeitpunkt pro Survey) und die zu testenden Studiendesigns (binäre Variable: Studientyp 1; Studientyp 2) eingingen. Mit der Variable Erhebungszeitpunkt wurde der Trend, mit der Variable Studiendesign der Methodeneffekt geschätzt. Um den Trendverlauf bestmöglich anhand der vorliegenden Daten zu schätzen,

wurden drei verschiedene Trendverlaufsmöglichkeiten geprüft: eine lineare Anpassung für einen geradlinigen Trend, eine quadratische Anpassung für einen U-förmigen oder abflachenden Trendverlauf und natürliche kubische Splines für komplexe und schwankende Trendverläufe (Annäherung des Trends durch stückweise Polynome 3. Grades, die an bestimmten Stellen aneinander gesetzt werden und die am Rand geradlinig verlaufen). Für jeden Indikator wurde die beste Anpassung an die Datenpunkte über das Bayes-Informationskriterium [40] ausgewählt. Bei Indikatoren, die nur zu drei Zeitpunkten vorlagen, wurde immer ein linearer Trendverlauf verwendet; ebenso in dem Fall, wenn über alle Zeitpunkte hinweg ein Studientyp nur einmalig vorlag. Die Modelle wurden sowohl für die Gesamtgruppe als auch stratifiziert nach Geschlechtsidentität, Alters- und Bildungsgruppen aufgestellt. Die Stratifizierung nach Bildung wurde in einer zusätzlichen Analyse eingeschränkt auf die Altersgruppe 30–64 Jahre, da sowohl bei jungen Erwachsenen (deren Bildungskarriere noch nicht abgeschlossen ist) als auch bei den 80-Jährigen und Älteren der Anteil mit niedriger Bildung erhöht ist.

Basierend auf den Ergebnissen der logistischen Regression lässt sich für jeden Zeitpunkt im Beobachtungszeitraum die Prävalenz des jeweiligen Indikators vorhersagen, wie sie für Studien vom Studientyp 1 bzw. vom Studientyp 2 erwartet wird. Daraus wurde der Methodeneffekt berechnet als Differenz zwischen den vom Modell vorhergesagten Prävalenzen für die beiden Studientypen:

$$\begin{aligned} \text{Methodeneffekt} = & \\ & \text{vorhergesagte Prävalenz für Studientyp 1} \\ & \text{minus} \\ & \text{vorhergesagte Prävalenz für Studientyp 2} \end{aligned}$$

Die Modelle beruhen auf der zentralen Annahme, dass der Methodeneffekt über die Zeit konstant ist. In der logistischen Regression wird der Methodeneffekt über das Odds Ratio (OR) modelliert, wobei das OR über die Zeit konstant ist. Da das OR eine schwierig zu interpretierende Maßzahl ist, berechnen wir den Methodeneffekt stattdessen, wie oben beschrieben, als Differenz. Diese Differenz kann für jeden Zeitpunkt bestimmt werden und variiert leicht über die Zeit, da die logistische Regression das OR und nicht die Differenz als konstant modelliert. Der Methodeneffekt wurde daher festgelegt als die Differenz zu dem Zeitpunkt, als der Methodenwechsel erfolgte (August 2024, dem mittleren Teilnahmezeitpunkt des RKI-Panels 2024).

Die beobachteten Prävalenzen aus den Studien vom Studientyp 2 (Telefonstichprobe mit telefonischem Interview) wurden in den Abbildungen zusätzlich um den Methodeneffekt korrigiert dargestellt, d. h. die Prävalenzen wurden an den Studientyp 1 (EMA-Stichprobe mit schriftlicher Be-

fragung) angeglichen, indem der Methodeneffekt zu den Prävalenzen vom Studientyp 2 addiert wurde. Analog wurden die Konfidenzintervallgrenzen umgerechnet.

Es gingen Personen ab 18 Jahren in die Auswertung ein. Für die Rekrutierungsstudie wurde die Datensatzversion v3 verwendet, für das RKI-Panel 2024 die Version v5. Die Trendauswertungen erfolgten gewichtet und für die EMA-Stichproben (Studientyp 1) unter Berücksichtigung der Clusterung in den Sample Points mit den Survey-Prozeduren von SAS 9.4 (SAS Institute, Cary, NC, USA). Die Berechnung der Methodeneffekte (Differenzen) mit 95 %-Konfidenzintervall erfolgte über das NLMeans-Makro in SAS. Jeder Survey wurde mit eigener Gewichtung berücksichtigt. Dabei wurden die Datensätze aus den verschiedenen Erhebungsjahren bezüglich der Alters- und Geschlechtsverteilung vergleichbar gemacht, indem die Gewichtung in jedem Survey an die europäische Standardbevölkerung [38] angepasst wurde. Für stratifizierte Auswertungen nach Bildung wurde die Standardisierung auch innerhalb der Bildungsgruppen vorgenommen, so dass in diesen Auswertungen allen Bildungsgruppen dieselbe Alters- und Geschlechtsverteilung zu Grunde liegt.

### 3. Ergebnisse

#### 3.1 Stichprobenzusammensetzung der Parallelerhebungen im Jahr 2024

##### Stichprobenzusammensetzung

Zum Vergleich der beiden Studientypen wird zunächst die Zusammensetzung der Stichproben in den beiden Parallelerhebungen verglichen, also der Telefonstudie GEDA 2024 und den ersten beiden Teilwellen (Q2 und Q3) des RKI-Panels 2024 (Tabelle 2). Der Blick auf die ungewichtete Stichprobenzusammensetzung zeigt, dass im RKI-Panel 2024 deutlich mehr junge Menschen bis 44 Jahre vertreten sind als in GEDA 2024. Diese Gruppe entspricht im RKI-Panel 2024 fast ihrem Bevölkerungsanteil. Auch Personen mit niedriger Bildung konnten im RKI-Panel 2024 besser erreicht werden, auch wenn sie immer noch unterrepräsentiert sind (19,3 % im Vergleich zu 33,4 % im Mikrozensus, in GEDA 2024 jedoch nur 14,2 %). Die bessere Abbildung der Bevölkerungsanteile nach Bildung zeigt sich dabei in allen Altersgruppen. Zudem stimmt der Anteil von Personen in Single-Haushalten im RKI-Panel 2024 (22,3 %) deutlich besser mit dem im Mikrozensus erhobenen Anteil (25,4 % der Erwachsenen) überein als in GEDA 2024 (34,1 %). Auch dieser Befund findet sich innerhalb der Altersgruppen wieder, wobei die Unterschiede mit steigendem Alter zunehmen. Der Anteil fehlender Werte in den Variablen zur Stichprobenzusammensetzung ist in beiden Studien vernachlässigbar (Höchstwert 0,4 %).

Die gewichteten Anteile nach Geschlecht und Alter stimmen in beiden Studien nahezu überein, da beide Studien auf

dieselbe Bevölkerungsverteilung gewichtet wurden. Unterschiede zeigen sich jedoch bei der gewichteten Bildungsverteilung und bei der Haushaltsgröße, weil in GEDA und im RKI-Panel nach unterschiedlichen Bildungsklassifikationen gewichtet wurde und weil die Haushaltsgröße bei der GEDA-Gewichtung nicht berücksichtigt wurde. Wird GEDA 2024 analog zum RKI-Panel gewichtet (Tabelle 2, Zeile 4), gleichen sich die gewichteten Anteile der Bildungsgruppen und der Personen in Single-Haushalten ebenfalls an.

#### 3.2 Prävalenzvergleich der Gesundheitsindikatoren in den Parallelerhebungen 2024

Tabelle 3 zeigt die Prävalenzen der Gesundheitsindikatoren in den Parallelerhebungen 2024. Hier werden nur gewichtete Prävalenzen dargestellt, sodass für Unterschiede in der Stichprobenzusammensetzung korrigiert wird, soweit dies durch die Gewichtung möglich ist. Für die nach Bildung stratifizierten Analysen werden aus diesem Grund die Prävalenzen mit Alters- und Geschlechtsstandardisierung innerhalb der Bildungsgruppen ausgegeben.

Die Prävalenzen der Indikatoren zum allgemeinen Gesundheitszustand (*Selbsteingeschätzte allgemeine Gesundheit* und *Chronisches Kranksein*) unterscheiden sich wenig zwischen den beiden Studientypen. Auffällig ist bei beiden Indikatoren die Altersgruppe der ab 80-Jährigen. Hier zeigt sich im RKI-Panel 2024 ein geringerer Anteil an Personen mit sehr guter oder guter allgemeiner Gesundheit (nicht signifikant) und ein deutlich höherer Anteil an chronisch kranken Personen im Vergleich zu GEDA 2024.

Besonders ausgeprägte Unterschiede zwischen den Parallelerhebungen zeigen die Prävalenzschätzer für die Indikatoren der psychischen Gesundheit. Die Prävalenz depressiver Symptomatik ist im RKI-Panel 2024 in der Gesamtgruppe und in fast allen Stratifizierungsgruppen höher als in GEDA 2024. Bei Frauen zeigen sich stärkere Unterschiede als bei Männern (ca. 8 % versus 3 % Differenz). Besonders groß sind die Unterschiede bei den jüngeren Altersgruppen (18- bis 29-Jährige ca. 16 % Differenz; 30- bis 44-Jährige ca. 6 %), während nur geringe oder keine Prävalenzunterschiede in den älteren Altersgruppen vorhanden sind. In allen Bildungsgruppen zeigt sich eine höhere Prävalenz im RKI-Panel 2024 als in GEDA 2024, wobei der Effekt in der niedrigen Bildungsgruppe stärker ausgeprägt ist, allerdings mit breiten Konfidenzintervallen in dieser Gruppe. Bei den Angstsymptomen liegen die Schätzer ebenfalls im RKI-Panel 2024 höher, dies ist aber weniger stark ausgeprägt. Die Unterschiede treten hauptsächlich bei Frauen (ca. 4 % Differenz) und 18- bis 29-Jährigen (ca. 11 % Differenz) auf. Sie finden sich vor allem bei der mittleren und hohen Bildungsgruppe (ca. 3 % Differenz).



**Tabelle 2:** Stichprobenzusammensetzung in den Parallelerhebungen Quartal 2 und Quartal 3 im Jahr 2024 (Zeilenprozente und 95 %-Konfidenzintervalle)

Survey und Studientyp		Geschlechtsidentität		Altersgruppe (Jahre)					Bildungsgruppe			Haushaltsgröße*	
		Weiblich	Männlich	18–29	30–44	45–64	65–79	≥ 80	Niedrig	Mittel	Hoch	Ein-Pers.- Haushalt	Mehr-Pers.- Haushalt
Anteile laut Statistischem Bundesamt*	%	51,1	48,9	16,0	23,5	33,7	18,1	8,7	33,4	46,6	20,0	25,4	74,6
Studientyp 1: RKI-Panel 2024 (Q2 + Q3) (n = 18.222)	n	9.853	8.324	2.641	3.894	6.035	4.150	1.502	3.519	8.706	5.965	4.062	14.157
	% ungewichtet	54,2 (53,5–54,9)	45,8 (45,1–46,5)	14,5 (13,9–15,1)	21,4 (20,6–22,2)	33,1 (32,3–33,9)	22,8 (22,1–23,5)	8,2 (7,8–8,7)	19,3 (18,4–20,3)	47,9 (46,8–49,0)	32,8 (31,3–34,4)	22,3 (21,4–23,2)	77,7 (76,8–78,6)
	% gewichtet	50,5 (49,6–51,5)	49,5 (48,5–50,4)	16,0 (15,2–16,7)	23,7 (22,9–24,5)	33,4 (32,6–34,2)	18,4 (17,9–19,0)	8,6 (8,1–9,1)	33,1 (31,7–34,6)	46,4 (45,2–47,7)	20,4 (19,2–21,7)	26,3 (25,3–27,4)	73,7 (72,6–74,7)
	Fehlende Werte (n, % gewichtet)	45 (0,2)		0 (0,0)					32 (0,2)			3 (0,0)	
Studientyp 2: GEDA 2024 (n = 4.016)	n	2.147	1.863	341	600	1.408	1.112	555	569	1.761	1.678	1.367	2.641
	% ungewichtet	53,5 (52,0–55,1)	46,5 (44,9–48,0)	8,5 (7,7–9,4)	14,9 (13,9–16,1)	35,1 (33,6–36,6)	27,7 (26,3–29,1)	13,8 (12,8–14,9)	14,2 (13,1–15,3)	43,9 (42,4–45,5)	41,9 (40,3–43,4)	34,1 (32,7–35,6)	65,9 (64,4–67,3)
	% gewichtet**	50,7 (48,3–53,2)	49,3 (46,8–51,7)	16,1 (14,1–18,3)	22,9 (20,7–25,4)	33,9 (31,6–36,2)	18,3 (16,7–20,0)	8,8 (7,7–10,1)	23,6 (21,3–26,0)	56,2 (53,8–58,6)	20,2 (18,8–21,6)	37,4 (34,9–39,9)	62,6 (60,1–65,1)
	% gewichtet analog RKI-Panel***	51,0 (48,6–53,4)	49,0 (46,6–51,4)	15,9 (14,1–18,0)	22,3 (20,1–24,5)	34,2 (32,1–36,5)	18,8 (17,3–20,5)	8,7 (7,8–9,8)	31,8 (29,3–34,4)	48,0 (45,7–50,4)	20,2 (18,8–21,6)	25,7 (23,7–27,7)	74,3 (72,3–76,3)
	Fehlende Werte (n, % gew.)	6 (0,2)		0 (0,0)					8 (0,4)			8 (0,3)	
p-Werte	p-Wert (ungew.)	0,44		<0,0001					<0,0001			<0,0001	
	p-Wert (gew.)**	0,87		0,97					<0,0001			<0,0001	
	p-Wert (gew.)***	0,72		0,75					0,47			0,56	

Die Angaben beziehen sich für das RKI-Panel 2024 auf Teilnehmende, die den Fragebogen C beantwortet haben, die Werte für Teilnehmende mit Fragebogen A sehen sehr ähnlich aus.

\*Im RKI-Panel 2024 wurde die Angabe zur Haushaltsgröße aus der Rekrutierungsstudie übernommen, falls sie in der Jahresehebung fehlte (z. B. wegen Nicht-Teilnahme in Quartal Q2).

\*Verteilung von Geschlecht und Alter: Fortschreibung der Bevölkerungszahlen zum 31.12.2023; Verteilung von Bildungsgruppe und Haushaltsgröße: Mikrozensus 2021

\*\*GEDA 2024 mit Gewichtung analog zu den vorangegangenen GEDA-Studien (Bildungsgewichtung mit ISCED, keine Gewichtung nach Haushaltsgröße)

\*\*\*GEDA 2024 mit Gewichtung analog zum RKI-Panel (Bildungsgewichtung mit CASMIN, mit Gewichtung nach Haushaltsgröße)

**Tabelle 3:** Vergleich der Gesundheitsindikatoren in den Parallelerhebungen in den Quartalen 2 und 3 im Jahr 2024 (gewichtete prozentuale Anteile und 95 %-Konfidenzintervalle)

Indikator (Anzahl gültige Werte in RKI-Panel 2024/ GEDA 2024)		Geschlechtsidentität			Altersgruppe (Jahre)					Bildungsgruppe		
		Gesamt	Weiblich	Männlich	18–29	30–44	45–64	65–79	≥80	Niedrig	Mittel	Hoch
Selbsteingeschätzte allgemeine Gesundheit Sehr gut/gut  (n = 18.519/4.010)	RKI-Panel 2024 (Q2 + Q3)	64,2 % (63,1–65,4)	61,5 % (60,0–63,0)	67,0 % (65,5–68,6)	78,3 % (75,7–80,7)	73,3 % (71,0–75,5)	63,2 % (61,4–64,9)	55,6 % (53,9–57,4)	35,2 % (31,9–38,5)	54,6 % (51,6–57,5)	67,5 % (66,1–68,8)	77,3 % (75,9–78,7)
	GEDA 2024	65,1 % (62,6–67,4)	65,4 % (62,0–68,6)	64,7 % (61,1–68,1)	82,9 % (76,8–87,7)	77,9 % (72,0–82,9)	60,6 % (56,5–64,5)	52,6 % (47,9–57,3)	42,0 % (35,1–49,2)	59,3 % (50,8–67,3)	65,4 % (61,9–68,7)	80,0 % (77,5–82,2)
	Fehlende Werte (n, % gewichtet)*	85 (0,6) 6 (0,3)	53 (0,8) 4 (0,2)	32 (0,4) 2 (0,5)	1 (0,2) 0 (0,0)	3 (0,2) 1 (0,1)	11 (0,2) 1 (0,6)	30 (0,8) 0 (0,0)	40 (3,2) 4 (1,0)	44 (0,9) 2 (0,1)	24 (0,3) 2 (0,4)	14 (0,3) 2 (0,1)
	p-Wert*	0,53	0,56		0,46					0,72		
	p-Wert innerhalb der Gruppen		0,0423	0,22	0,1597	0,1443	0,24	0,24	0,0811	0,30	0,27	0,0638
Chronisches Kranksein Ja  (n = 18.513/4.006)	RKI-Panel 2024 (Q2 + Q3)	54,3 % (53,2–55,4)	58,1 % (56,8–59,5)	50,2 % (48,6–51,8)	33,9 % (31,3–36,5)	42,4 % (40,3–44,6)	60,5 % (59,0–62,1)	67,2 % (65,4–69,0)	73,3 % (70,2–76,3)	57,5 % (54,9–60,0)	52,6 % (51,2–54,0)	48,6 % (47,0–50,2)
	GEDA 2024	53,2 % (50,7–55,7)	55,7 % (52,2–59,1)	50,8 % (47,2–54,4)	32,5 % (26,1–39,7)	43,6 % (37,6–49,7)	60,0 % (56,0–63,7)	67,3 % (63,1–71,3)	60,9 % (53,8–67,5)	47,0 % (38,4–55,7)	55,3 % (51,7–58,8)	43,5 % (40,3–46,8)
	Fehlende Werte (n, % gewichtet)*	91 (0,7) 10 (0,1)	57 (0,9) 5 (0,1)	34 (0,4) 5 (0,1)	4 (0,2) 0 (0,0)	5 (0,3) 1 (0,1)	18 (0,5) 2 (0,1)	31 (0,8) 6 (0,3)	33 (3,0) 1 (0,0)	41 (1,0) 1 (0,0)	31 (0,4) 3 (0,1)	16 (0,4) 6 (0,3)
	p-Wert*	0,44	0,50		0,36					0,21		
	p-Wert innerhalb der Gruppen		0,1900	0,76	0,72	0,73	0,78	0,97	0,0006	0,0231	0,1748	0,0069
Adipositas BMI ≥ 30  (n = 18.450/3.918)	RKI-Panel 2024 (Q2 + Q3)	22,4 % (21,4–23,4)	21,6 % (20,4–22,9)	23,2 % (21,9–24,6)	14,8 % (12,5–17,5)	22,7 % (20,8–24,8)	25,8 % (24,3–27,4)	24,9 % (23,2–26,7)	17,4 % (15,0–20,1)	30,0 % (27,4–32,8)	21,9 % (20,8–23,0)	12,9 % (11,8–14,0)
	GEDA 2024	19,4 % (17,5–21,5)	17,4 % (14,9–20,3)	21,5 % (18,5–24,7)	8,2 % (5,0–13,1)	19,1 % (14,5–24,8)	23,2 % (19,9–26,8)	24,3 % (20,2–28,8)	15,8 % (11,1–22,0)	30,3 % (22,2–40,0)	18,7 % (16,0–21,6)	12,7 % (10,7–14,9)
	Fehlende Werte (n, % gewichtet)*	154 (1,1) 98 (2,7)	98 (1,4) 74 (4,2)	55 (0,8) 23 (1,1)	20 (0,9) 12 (3,2)	16 (0,6) 13 (2,8)	44 (1,1) 33 (2,1)	39 (1,0) 17 (1,5)	35 (3,0) 23 (6,6)	59 (1,8) 18 (4,8)	62 (0,8) 48 (2,7)	33 (0,5) 32 (1,6)
	p-Wert*	0,0119	0,0126		0,0097					0,24		
	p-Wert innerhalb der Gruppen		0,0081	0,31	0,0179	0,22	0,1839	0,78	0,61	0,95	0,0448	0,89
Rauchen Ja, täglich/ ja, gelegentlich  (n = 18.318/4.014)	RKI-Panel 2024** (Q2 + Q3)	23,7 % (22,7–24,6)	20,4 % (19,2–21,7)	27,0 % (25,6–28,4)	27,1 % (24,6–29,7)	30,9 % (28,7–33,2)	26,3 % (24,7–28,0)	15,3 % (13,8–16,9)	3,7 % (2,6–5,1)	36,6 % (33,7–39,6)	23,4 % (22,3–24,5)	13,6 % (12,3–14,9)
	GEDA 2024	27,0 % (24,7–29,5)	21,9 % (18,9–25,2)	32,3 % (28,9–36,0)	29,2 % (23,0–36,4)	35,5 % (29,8–41,7)	30,1 % (26,3–34,3)	19,5 % (15,9–23,7)	4,5 % (2,4–8,4)	38,7 % (30,3–47,8)	28,3 % (25,1–31,8)	16,0 % (13,5–18,8)
	Fehlende Werte (n, % gewichtet)*	286 (1,7) 2 (0,0)	187 (2,3) 2 (0,1)	99 (1,1) 0 (0,0)	18 (0,7) 0 (0,0)	19 (0,9) 0 (0,0)	36 (0,8) 1 (0,1)	106 (2,6) 1 (0,0)	107 (7,7) 0 (0,0)	133 (2,6) 0 (0,0)	89 (1,1) 2 (0,0)	57 (1,0) 0 (0,0)
	p-Wert*	0,0088	0,0080		0,0050					0,0132		
	p-Wert innerhalb der Gruppen		0,39	0,0046	0,55	0,1488	0,0761	0,0386	0,58	0,65	0,0042	0,0968

\*Anzahl und Anteil fehlender Werte: Die erste Zeile steht für das RKI-Panel 2024, die zweite Zeile für GEDA 2024.

\*\*Im RKI-Panel 2024 wurde der Rauchstatus von Personen mit Diabetes mellitus aus der Rekrutierungsstudie übernommen.

\*p-Wert für Unterschied zwischen RKI-Panel 2024 (Q2 + Q3) und GEDA 2024 im Gesamt (ohne Adjustierung), bzw. im Gesamt adjustiert für Geschlecht/Altersgruppe/Bildung

BMI = Body-Mass-Index, PHQ-8 = Patient Health Questionnaire-8, GAD-2 = Generalized Anxiety Disorder-2

Analyse bezogen auf Bevölkerungsstand 31.12.2023, für die Stratifizierung nach Bildung innerhalb der Bildungsgruppen standardisiert nach europäischer Standardbevölkerung (ESP 2013)

▼ Fortsetzung nächste Seite ▼

**Tabelle 3 Fortsetzung:** Vergleich der Gesundheitsindikatoren in den Parallelerhebungen in den Quartalen 2 und 3 im Jahr 2024 (gewichtete prozentuale Anteile und 95 %-Konfidenzintervalle)

Indikator (Anzahl gültige Werte in RKI-Panel 2024/ GEDA 2024)		Geschlechtsidentität			Altersgruppe (Jahre)					Bildungsgruppe		
		Gesamt	Weiblich	Männlich	18–29	30–44	45–64	65–79	≥80	Niedrig	Mittel	Hoch
Selbstberichtete psychische Gesundheit Ausgezeichnet/sehr gut  (n = 18.042/4.004)	RKI-Panel 2024 (Q2 + Q3)	44,9 % (43,9–45,9)	40,4 % (39,1–41,8)	49,6 % (48,1–51,0)	41,8 % (39,0–44,6)	45,6 % (43,5–47,8)	47,0 % (45,3–48,7)	48,3 % (46,4–50,2)	32,9 % (29,8–36,3)	38,0 % (35,3–40,7)	45,8 % (44,4–47,2)	54,2 % (52,5–55,8)
	GEDA 2024	39,6 % (37,2–42,0)	35,4 % (32,3–38,7)	43,8 % (40,3–47,3)	45,9 % (38,7–53,2)	45,3 % (39,5–51,2)	40,6 % (36,9–44,4)	36,6 % (32,5–40,9)	15,2 % (11,7–19,4)	29,7 % (21,9–38,9)	39,2 % (35,8–42,7)	53,6 % (50,3–56,9)
	Fehlende Werte (n, % gewichtet) <sup>†</sup>	180 (1,2)	96 (1,1)	84 (1,2)	57 (2,3)	39 (1,0)	34 (0,7)	21 (0,8)	29 (1,9)	48 (1,5)	86 (1,1)	46 (0,9)
		12 (0,3)	8 (0,3)	4 (0,4)	0 (0,0)	0 (0,0)	3 (0,1)	5 (0,6)	4 (2,2)	6 (0,5)	4 (0,1)	2 (0,1)
	p-Wert*	<0,0001	<0,0001		<0,0001					0,0003		
	p-Wert innerhalb der Gruppen		0,0053	0,0033	0,30	0,91	0,0028	<0,0001	<0,0001	0,0868	0,0006	0,78
Depressive Symptomatik PHQ-8 ≥ 10  (n = 17.856/3.841)	RKI-Panel 2024 (Q2 + Q3)	20,7 % (19,8–21,6)	23,2 % (22,0–24,5)	17,8 % (16,6–19,1)	32,2 % (29,6–34,9)	22,7 % (20,8–24,8)	19,7 % (18,3–21,1)	11,1 % (9,8–12,5)	17,1 % (14,2–20,4)	28,8 % (25,9–31,8)	19,7 % (18,6–20,8)	14,3 % (13,2–15,6)
	GEDA 2024	15,1 % (13,1–17,3)	15,6 % (12,8–18,7)	14,7 % (11,9–18,0)	15,9 % (11,3–21,7)	16,4 % (11,9–22,3)	16,7 % (13,4–20,6)	8,2 % (5,3–12,4)	18,2 % (12,2–26,3)	18,7 % (12,2–27,5)	16,1 % (13,3–19,4)	9,2 % (7,2–11,7)
	Fehlende Werte (n, % gewichtet) <sup>†</sup>	366 (2,4)	207 (2,7)	159 (2,0)	39 (1,9)	23 (0,6)	56 (1,3)	135 (4,1)	113 (8,9)	159 (3,1)	135 (1,7)	65 (1,3)
		175 (4,4)	108 (4,3)	67 (4,4)	7 (1,9)	11 (3,0)	33 (2,9)	71 (7,8)	53 (11,1)	52 (4,7)	62 (3,3)	60 (2,9)
	p-Wert*	<0,0001	<0,0001		<0,0001					0,0004		
	p-Wert innerhalb der Gruppen		<0,0001	0,0768	<0,0001	0,0443	0,1477	0,1687	0,78	0,0317	0,0428	0,0005
Angstsymptome GAD-2 ≥ 3  (n = 18.039/3.970)	RKI-Panel 2024 (Q2 + Q3)	15,4 % (14,7–16,3)	18,4 % (17,3–19,5)	12,3 % (11,3–13,3)	27,3 % (24,7–30,1)	17,8 % (16,2–19,6)	13,8 % (12,7–15,1)	7,6 % (6,6–8,7)	9,8 % (7,7–12,4)	17,7 % (15,3–20,3)	15,9 % (15,0–17,0)	13,7 % (12,5–15,0)
	GEDA 2024	13,2 % (11,3–15,3)	14,6 % (11,9–17,7)	11,8 % (9,3–14,8)	16,2 % (11,9–21,8)	17,9 % (13,0–24,0)	13,4 % (10,5–17,1)	5,4 % (3,5–8,2)	10,7 % (6,1–18,1)	16,8 % (10,6–25,6)	13,2 % (10,6–16,4)	10,9 % (8,6–13,7)
	Fehlende Werte (n, % gewichtet) <sup>†</sup>	183 (1,3)	103 (1,4)	80 (1,1)	42 (1,7)	26 (0,7)	28 (0,7)	42 (1,3)	45 (4,0)	71 (1,8)	71 (0,8)	39 (0,9)
		46 (1,5)	32 (1,9)	14 (1,0)	2 (1,5)	5 (0,9)	8 (0,4)	11 (0,7)	20 (8,7)	14 (2,0)	16 (0,9)	16 (0,8)
	p-Wert*	0,0502	0,0619		0,0500					0,1003		
	p-Wert innerhalb der Gruppen		0,0260	0,76	0,0007	0,99	0,82	0,1198	0,79	0,83	0,1024	0,0646

<sup>†</sup>Anzahl und Anteil fehlender Werte: Die erste Zeile steht für das RKI-Panel 2024, die zweite Zeile für GEDA 2024.

<sup>\*\*</sup>Im RKI-Panel 2024 wurde der Rauchstatus von Personen mit Diabetes mellitus aus der Rekrutierungsstudie übernommen.

\*p-Wert für Unterschied zwischen RKI-Panel 2024 (Q2 + Q3) und GEDA 2024 im Gesamt (ohne Adjustierung), bzw. im Gesamt adjustiert für Geschlecht/Altersgruppe/Bildung

BMI = Body-Mass-Index, PHQ-8 = Patient Health Questionnaire-8, GAD-2 = Generalized Anxiety Disorder-2

Analyse bezogen auf Bevölkerungsstand 31.12.2023, für die Stratifizierung nach Bildung innerhalb der Bildungsgruppen standardisiert nach europäischer Standardbevölkerung (ESP 2013)

Beim Indikator der selbsteingeschätzten psychischen Gesundheit zeigte sich im RKI-Panel 2024 ein höherer Anteil an Personen mit ausgezeichneter oder sehr guter psychischer Gesundheit, verglichen mit GEDA 2024. Da dieses Ergebnis im Widerspruch zu den gerade beschriebenen Befunden zur psychischen Gesundheit steht und eine Anpassung der Analyse nach sich zog, soll hier bereits eine mögliche Erklärung aufgezeigt werden. Vergleicht man die Prävalenzen aus der Jahrerhebung des RKI-Panels 2024 mit den entsprechenden Prävalenzen aus der nur wenige Monate zuvor durchgeführten Rekrutierungsstudie, wird in der Rekrutierungsstudie die eigene psychische Gesundheit (von denselben Personen) deutlich schlechter eingeschätzt als in der Jahrerhebung. Selbst wenn man auf Personen einschränkt, die bereits in der ersten Teilwelle der Jahrerhebung (Quartal 2) den Fragebogen C ausgefüllt haben und damit in kurzem Abstand zur Rekrutierungsstudie von durchschnittlich 83 Tagen die gleiche Frage erneut gestellt bekamen, liegt der Unterschied im Anteil von ausgezeichneter/sehr guter psychischer Gesundheit bei 45 % im RKI-Panel 2024 versus 34 % in der Panel-Rekrutierungsstudie. Als mögliche Erklärung für diese großen Differenzen wurde die Positionierung der Frage im Fragebogen identifiziert. In der kurzen Rekrutierungsstudie mit insgesamt 26 Fragen wurde diese Frage an Position 16 nach Fragen zur allgemeinen Gesundheit sowie zum Gesundheitsverhalten gestellt. Im Fragebogen C im RKI-Panel 2024 erschien die Frage weiter hinten, an Position 24, eingeordnet im Anschluss an Fragen zu depressiver Symptomatik, Fatigue (Erschöpfung) und suizidalen Gedanken. Im Kontext dieser Fragen haben die Teilnehmenden ihre eigene Situation offenbar weniger schlecht eingeschätzt. Aufgrund dieses Kontext- bzw. Fragebogeneffekts sind die Angaben aus der Jahrerhebung des RKI-Panels 2024 nicht mit bisherigen Ergebnissen der GEDA-Studien vergleichbar, in denen diese Frage den Abschnitt zur psychischen Gesundheit eingeleitet hatte. Der Effekt verstärkt sich offenbar mit dem Alter, da die Abweichungen zu den Prävalenzen aus der Rekrutierungsstudie mit zunehmendem Alter größer ausfallen. In den im folgenden Abschnitt beschriebenen Trendanalysen werden daher für diesen Indikator die Angaben aus der Rekrutierungsstudie des Panels anstelle der Jahrerhebung berücksichtigt.

Die Adipositasprävalenz liegt im RKI-Panel 2024 um drei Prozentpunkte höher als in GEDA 2024. Der Effekt ist etwas stärker bei Frauen und auffallend deutlicher bei den 18- bis 29-Jährigen (mehr als 6 Prozentpunkte Unterschied). Die Rauchprävalenz wiederum ist im RKI-Panel 2024 über drei Prozentpunkte niedriger als in GEDA 2024, mit deutlicheren Unterschieden bei Männern, den mittleren Altersgruppen und der mittleren Bildungsgruppe.

Die in Sensitivitätsanalysen durchgeführte Gewichtung von GEDA 2024 analog zum RKI-Panel wirkte sich kaum auf die Prävalenzschätzer aus. Die Änderung lag meist zwischen

null bis maximal zwei Prozentpunkten und verstärkte tendenziell die Größe der Methodeneffekte. Nur beim Indikator Rauchen verringerte sich durch die veränderte Gewichtung der Methodeneffekt um etwa zwei Prozentpunkte und verschwand bei den Frauen ganz.

### 3.3 Antwortverhalten in Abhängigkeit vom Erhebungsmodus in den Parallelerhebungen 2024

Dieser Abschnitt beschreibt Muster im Antwortverhalten, die sich durch den veränderten Erhebungsmodus bei verschiedenen Frage- und Antwortformaten ergeben, wie kategorialem oder metrischen Antworten bzw. einfachen oder mehrfachen Likert-Skalen. Darüber hinaus werden die Anteile fehlender Werte in Abhängigkeit vom Antwortformat beschrieben.

#### Antwortformat: Einfache Likert-Skala

Die beiden Indikatoren *Selbsteingeschätzte allgemeine Gesundheit* und *Selbsteingeschätzte psychische Gesundheit* basieren jeweils auf einer einzelnen Likert-Skala mit fünf Antwortmöglichkeiten, deren zwei erste Antwortmöglichkeiten für den jeweiligen Indikator zusammengefasst werden.

Der Indikator der *Selbsteingeschätzten allgemeinen Gesundheit* („sehr gut“ oder „gut“) zeigt im Parallelvergleich nur geringe Unterschiede, wie bereits im vorigen Abschnitt beschrieben (64 % im RKI-Panel 2024 vs. 65 % in GEDA 2024, [Tabelle 3](#)). Eine Betrachtung der Einzelkategorien offenbart jedoch Verschiebungen: In GEDA 2024 geben mehr Personen „sehr gut“ (19 %) an als im RKI-Panel 2024 (12 %), während „gut“ im RKI-Panel häufiger genannt wird (52 % vs. 46 %). Auch in Bezug auf die Einzelkategorien am anderen Ende der Skala werden in der schriftlichen Befragung im RKI-Panel Randkategorien seltener und mittlere Kategorien häufiger gewählt.

Wegen der bereits oben beschriebenen widersprüchlichen Ergebnisse für den Indikator der selbsteingeschätzten psychischen Gesundheit, die wahrscheinlich auf Reihenfolgeeffekte im Fragebogen zurückzuführen sind, wird das Antwortverhalten für diesen Indikator für das RKI-Panel anhand der identischen Frage aus der Rekrutierungsstudie analysiert. Dort zeigt sich – analog zur selbsteingeschätzten allgemeinen Gesundheit – kaum ein Unterschied bei der Betrachtung des Indikators mit den kombinierten Kategorien („ausgezeichnet“ oder „sehr gut“), mit 25 % in GEDA 2024 vs. 26 % in der Rekrutierungsstudie des RKI-Panels 2024. Bei der Betrachtung der Einzelkategorien zeigt sich im RKI-Panel 2024 wieder das in schriftlichen Befragungen erwartbare Antwortmuster mit selteneren Rand- und häufigeren Mittelkategorien: Die Kategorie „ausgezeichnet“ wurde schriftlich seltener gewählt (8,1 % vs. 11,7 %), während „sehr gut“ häufiger angegeben wurde (17,9 % vs. 13,2 %).



Antwortformat: Mehrfache Likert-Skalen

Die Indikatoren *Depressive Symptomatik* und *Angstsymptome* beruhen auf acht bzw. zwei Items mit Likert-Skalen-Antwortformat, denen Punktwerte zugeordnet sind. Diese werden zu einem Summenwert addiert und anhand eines festgesetzten Schwellenwerts klassifiziert.

Besonders deutliche Unterschiede zwischen GEDA 2024 und dem RKI-Panel 2024 zeigen sich beim Indikator für depressive Symptomatik (siehe oben beim Prävalenzvergleich, [Tabelle 3](#)), der auf acht Einzelitems (PHQ1 bis PHQ8) basiert. In allen Einzelitems ist im RKI-Panel im Vergleich zur telefonischen Erhebung eine Verschiebung vom Rand zu den mittleren Antwortmöglichkeiten zu beobachten ([Abbildung 1](#)). Die Randkategorien „Überhaupt nicht“ und „Beinahe jeden Tag“ (Grautöne in der Abbildung) werden im RKI-Panel seltener gewählt, während die mittleren Kategorien „An einzelnen Tagen“ und „An mehr als der Hälfte der Tage“ häufiger angegeben werden (Blautöne in der Abbildung). Da die Kategorie „Überhaupt nicht“ in beiden Erhebungen deutlich häufiger genannt wird als „Beinahe jeden Tag“, führt die beobachtete Verschiebung im RKI-Panel 2024 insgesamt zu höheren Summenwerten. In der Folge fällt die Prävalenz depressiver Symptomatik im RKI-Panel 2024 höher aus als in GEDA 2024.

Beim auf zwei Items (GAD1, GAD2) basierenden Indikator für Angstsymptome zeigt sich das gleiche Bild wie

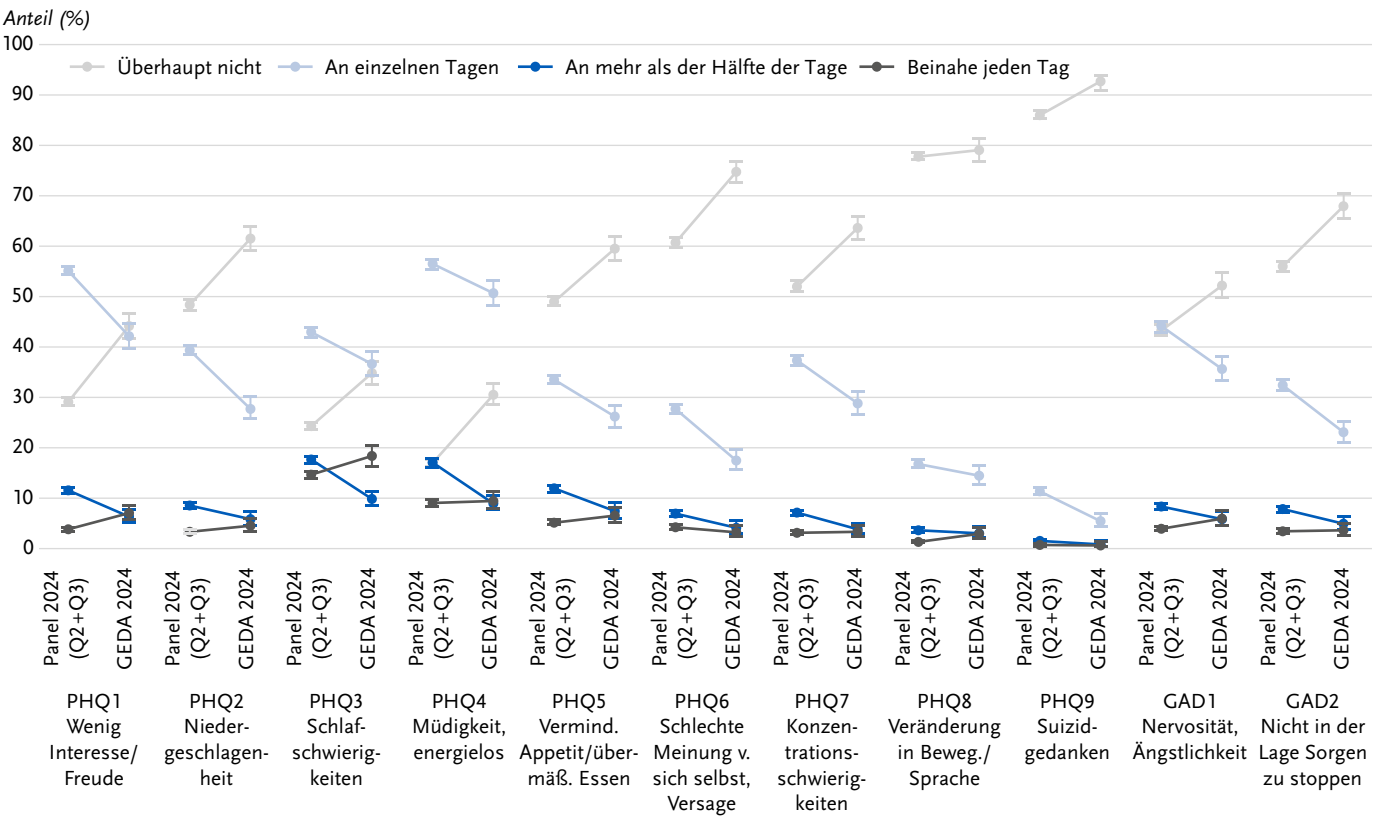
bei den acht Einzelitems der depressiven Symptomatik ([Abbildung 1](#)). Auch hier werden in der schriftlichen Befragung die mittleren Kategorien häufiger angegeben als am Telefon. Da eine Aufsummierung nur über zwei Items stattfindet, ist die Auswirkung auf die Prävalenzschätzer bei den Angstsymptomen geringer als bei der depressiven Symptomatik ([Tabelle 3](#)).

Antwortformat: Kategorisierung von metrischen Angaben

Der Indikator *Adipositas* beruht auf jeweils metrischen, selbstberichteten Angaben zu Körpergröße und Gewicht, aus denen der BMI gebildet wird. Die Angaben zur Körpergröße unterscheiden sich in beiden Erhebungen nicht. Die Angaben zum Körpergewicht sind im Panel 2024 im Mittel etwas höher als in GEDA 2024 (79,9 kg versus 78,3 kg). Auch die Perzentile liegen im Panel 2024 höher als in GEDA 2024 (z. B. 90 %-Perzentil 104 kg vs. 100 kg oder 95 %-Perzentil 115 kg vs. 110 kg). Aus diesen relativ geringen Änderungen am mittleren Gewicht, die zu höheren BMI-Werten führen, ergeben sich durch die feste Kategorisierung des BMI  $\geq 30$  kg/m<sup>2</sup> nicht vernachlässigbare Prävalenzunterschiede für Adipositas in Höhe von 3 Prozentpunkten.

Antwortformat: Kategorial

Beim Indikator *Chronisches Kranksein* als einzigem Indikator, der auf einer einfachen Ja/Nein-Frage basiert, zeigen sich



**Abbildung 1:** Verschiebung der Anteile der einzelnen Antwortkategorien für die acht Einzelitems des Indikators Depressive Symptomatik (PHQ-8) und die zwei Einzelitems des Indikators Angstsymptome (GAD-2) in den Parallelerhebungen in den Quartalen 2 und 3 im Jahr 2024 (gewichtete Analysen), graue Kategorien sind Randkategorien, blaue Kategorien sind Mittelkategorien. Quelle: Parallelerhebungen GEDA 2024 und RKI-Panel 2024 (Q2 + Q3) PHQ-8 = Patient Health Questionnaire-8, GAD-2 = Generalized Anxiety Disorder-2

keine Unterschiede im Antwortverhalten zwischen den Erhebungsmodi (Tabelle 3).

Der Indikator *Rauchen* basiert auf einer kategorialen Abfrage mit vier Antwortmöglichkeiten. Die im Panel 2024 im Vergleich GEDA 2024 beobachtete niedrigere Prävalenz aktuell Rauchender („ja, täglich“ oder „ja, gelegentlich“) ist vor allem auf einen geringeren Anteil der erstgenannten Kategorie („ja, täglich“) im Panel 2024 zurückzuführen (18,0 % vs. 20,9 %) sowie auf einen höheren Anteil von Personen, die angaben, noch nie geraucht zu haben (46,6 % vs. 44,2 %), welches die letztgenannte Kategorie darstellt. Die mittleren Kategorien (gelegentlich oder ehemals Rauchende) weisen in den beiden Parallelerhebungen ähnliche Prävalenzen auf. Eine für telefonische Befragungen typische Verschiebung hin zu den Randkategorien, da Erst- und Letztgehörtes häufiger widergegeben wird, ist dabei nicht erkennbar.

#### Anteil fehlender Werte in Abhängigkeit vom Antwortformat

Die Anteile fehlender Werte sind in Tabelle 3 enthalten. Generell liegen die Anteile in GEDA 2024 sehr niedrig und sind im RKI-Panel 2024 etwas höher, was vor allem auf den Einsatz von Papierfragebögen im Panel zurückzuführen ist. Aber auch im Online-Modus zeigen sich leicht höhere Anteile von fehlenden Werten im Vergleich zum Telefoninterview. Dies gilt etwa für die beiden Indikatoren zum allgemeinen Gesundheitszustand (*Selbsteingeschätzte allgemeine Gesundheit, Chronisches Kranksein*), bei denen eine einfache Likert-Skala bzw. eine kategoriale Variable vorliegt. Hier liegen die Anteile fehlender Werte in allen Subgruppen und beiden Studien bei maximal 1 %, meist deutlich darunter. Eine Ausnahme bilden die Hochaltrigen ab 80 Jahren, die im RKI-Panel bei beiden Indikatoren rund 3 % fehlende Werte aufweisen (im Vergleich zu 1 % bzw. 0 % in GEDA 2024), was auf den hohen Anteil an Papierfragebögen in dieser Altersgruppe zurückzuführen ist. Beim kategorialen Indikator *Rauchen* finden sich fast keine fehlenden Werte im Telefoninterview in GEDA 2024, jedoch ein Anteil von 1,7 % in der Gesamtgruppe im RKI-Panel 2024, der wiederum in den oberen Altersgruppen und bei niedriger Bildung etwas höher liegt (bis zu 8 % bzw. knapp 3 %).

Bei mehrfachen Likert-Skalen (*Depressive Symptomatik, Angstsymptome*) nimmt der Anteil fehlender Werte mit der Anzahl der zugrundeliegenden Fragen zu, da die Summenwerte nur berechnet werden, wenn alle Einzelfragen beantwortet wurden. Bei den Angstsymptomen (2 Items) liegt der Anteil fehlender Werte meist im Bereich von 1 bis 2 %, wobei keine klaren Unterschiede zwischen den Studientypen erkennbar sind. Bei der depressiven Symptomatik (8 Items) treten im RKI-Panel in der Gesamtgruppe 2,4 % fehlende Werte auf, in GEDA 2024 liegt der Anteil mit 4,4 % deutlich höher. Dieser Unterschied zwischen den Studientypen findet sich in allen Subgruppen. In beiden Studientypen und bei beiden Indikatoren liegt der Anteil fehlender Werte bei den

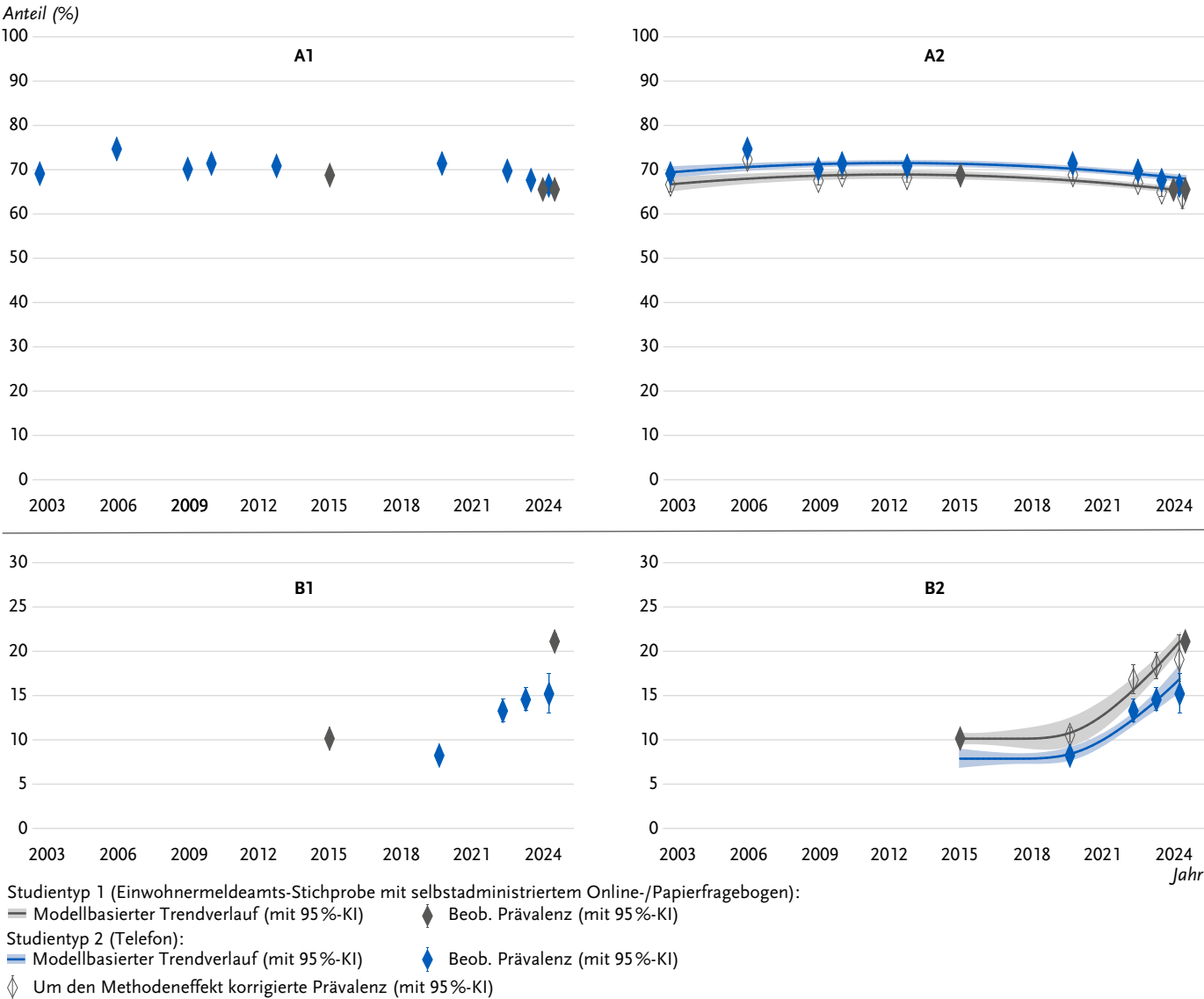
höheren Altersgruppen und in der niedrigen Bildungsgruppe teilweise deutlich höher als in den jeweiligen Vergleichsgruppen.

Höhere Anteile fehlender Werte im Telefoninterview in GEDA 2024 (2,7 %) im Vergleich zum RKI-Panel 2024 (1,1 %) finden sich auch bei der Adipositasprävalenz. Der Unterschied zwischen den Studien ist am stärksten ausgeprägt bei Frauen (4,2 % in GEDA 2024 vs. 1,4 % im RKI-Panel 2024) und jungen Erwachsenen (3,2 % vs. 0,9 %).

#### 3.4 Methodeneffekt im Zeitverlauf

Die nun folgende alters- und geschlechtsstandardisierte Auswertung bezieht die komplette Zeitreihe der RKI-Befragungssurveys von 2003 bis 2024 ein. Sie berücksichtigt neben der Jahrerhebung des RKI-Panels 2024 (hier nun ohne Einschränkung auf Quartal 2 und 3) auch die Daten aus der Rekrutierungsstudie. Ziel ist die Beurteilung des Prävalenzverlaufs über die Zeit, bei Kontrolle für den Methodeneffekt.

Abbildung 2 zeigt das Ergebnis der modellbasierten Zeitreihenauswertung für zwei exemplarische Indikatoren. Die kompletten Zeitreihenauswertungen einschließlich Stratifizierung sind für die verschiedenen Indikatoren in Indikatorenblättern dargestellt, die auf dem Publikationsserver des RKI (<https://edoc.rki.de/>) veröffentlicht werden. Beispiel A (*Selbsteingeschätzte allgemeine Gesundheit*) zeigt einen Indikator mit einer großen Zahl von Datenpunkten, davon drei vom Studientyp 1, einen in der Mitte und zwei am Ende der Beobachtungszeit, was eine relativ belastbare Schätzung des Methodeneffekts ermöglicht. Der Anteil der Erwachsenen mit einer sehr guten oder guten allgemeinen Gesundheit ist seit 2003 bei den Telefonstichproben leicht rückläufig, insbesondere in der Zeit nach der COVID-19-Pandemie (von 69,0 % im GSTel03 auf 69,6 % in GEDA 2022 und auf 67,5 % in GEDA 2023). Bei den EMA-Stichproben zeigt sich ebenfalls ein rückläufiger Trend auf einem etwas niedrigeren Niveau, von 68,7 % in GEDA 2014/2015-EHIS auf 65,4 % im Panel 2024. Der Methodeneffekt, also der Prävalenzunterschied zwischen den beiden Studientypen, ist mit einer Differenz von -2,8 Prozentpunkten moderat und zeigt niedrigere Prävalenzen in den EMA-Stichproben an. Der vom Modell geschätzte parallele Trendverlauf für die beiden Studientypen passt sich gut an die Prävalenzschätzer aus den einzelnen Studien an, was die Annahme eines konstanten Methodeneffekts über die Zeit gerechtfertigt erscheinen lässt. Beurteilt man den Zeitverlauf naiv durch den direkten Vergleich von GEDA 2023 mit Panel 2024, ergibt sich eine Abnahme der Prävalenz um 2,1 Prozentpunkte. Betrachtet man dagegen den für den Methodeneffekt korrigierten Verlauf der Trendlinien vom Jahr 2023 zum Jahr 2024, wird der rückläufige Trend auf nur 0,6 Prozentpunkte geschätzt, d. h. der zeitliche Trend würde ohne Modellierung des Methodeneffekts überschätzt.



**Abbildung 2:** Darstellung der Prävalenzentwicklung über die Zeit (2003–2024) unter Berücksichtigung des Methodeneffekts für zwei exemplarische Indikatoren (Analyse für Gesamt ohne Stratifizierung). Quelle: RKI-Gesundheitssurveys (GSTel, GEDA), Rekrutierungsstudie für das RKI-Panel 2024, GEDA 2024, RKI-Panel 2024.  
A: Indikator Selbsteingeschätzte allgemeine Gesundheit, B: Indikator Depressive Symptomatik  
A1 bzw. B1: Prävalenzschätzer aus den verschiedenen Surveys  
A2 bzw. B2: Zusätzlich geschätzter Trendverlauf für Studientyp 1 bzw. Studientyp 2

In Beispiel B (*Depressive Symptomatik*) liegen weniger Datenpunkte vor, aber auch hier liegen zwei Surveys vom Studientyp 1 jeweils an den beiden Rändern des Beobachtungszeitraums. Die Prävalenzen zeigen einen zunehmenden Anteil an depressiver Symptomatik seit GEDA 2014/2015-EHIS, auch hier mit einer deutlichen Zunahme nach der Pandemie, die von dem flexiblen Spline-Modell nicht optimal abgebildet werden kann. Dennoch scheint offensichtlich, dass es einen relativ großen Methodeneffekt von etwa 4 Prozentpunkten gibt. Im direkten Vergleich von GEDA 2023 mit 14,6 % zu Panel 2024 mit 21,2 % zeigt sich ein Trendanstieg von fast 7 Prozentpunkten, während das Modell mit Korrektur für den Methodeneffekt einen Anstieg von nur 2 bis 3 Prozentpunkten vorhersagt.

[Tabelle 4](#) gibt eine Übersicht über die geschätzten Methodeneffekte, ausgedrückt über die Differenz der beiden

Trendlinien. Es zeigen sich stärkere Methodeneffekte bei den Indikatoren der psychischen Gesundheit, der sozialen Unterstützung und der Rauchprävalenz im Vergleich zu den Indikatoren der körperlichen Gesundheit und dem Indikator Adipositas.

Die Methodeneffekte über die Zeit sind bei den Indikatoren der psychischen Gesundheit bei Frauen stärker ausgeprägt als bei Männern und nehmen mit dem Alter ab (Ausnahme: Methodeneffekt der Altersgruppe ab 80 Jahre für den Indikator Angstsymptome). Insgesamt zeigen sich besonders starke Effekte für die Altersgruppe der 18- bis 29-Jährigen. So wird beispielsweise beim Indikator Depressive Symptomatik der Effekt auf die Prävalenzänderung, der sich aufgrund des Methodenwechsels erklären lässt, in dieser Altersgruppe auf 11 Prozentpunkte geschätzt.

**Tabelle 4:** Geschätzter Methodeneffekt auf die Prävalenzen über die Zeit im Vergleich von Studientyp 1 zu Studientyp 2 (Prävalenzdifferenz und 95%-Konfidenzintervalle)

	Geschlechtsidentität			Altersgruppe (Jahre)					Bildungsgruppe		
	Gesamt	Weiblich	Männlich	18–29	30–44	45–64	65–79	≥ 80	Niedrig	Mittel	Hoch
Selbsteingeschätzte allgemeine Gesundheit (Sehr gut/gut)	-2,8 (-3,4--2,1)	-3,1 (-4,0--2,2)	-2,6 (-3,4--1,7)	-3,9 (-5,4--2,4)	-3,0 (-4,3--1,7)	-0,6 (-1,7--0,4)	-3,1 (-4,4--1,9)	-10,9 (-13,0--8,7)	-3,0 (-4,7--1,2)	-2,2 (-3,0--1,3)	-2,1 (-3,0--1,2)
Chronisches Kranksein (ja)	2,8 (2,1--3,5)	2,9 (2,0--3,8)	2,9 (1,9--3,8)	1,9 (0,2--3,7)	0,3 (-1,2--1,8)	2,7 (1,7--3,8)	4,2 (3,0--5,4)	12,2 (10,2--14,1)	2,7 (0,9--4,5)	2,7 (1,8--3,6)	3,8 (2,8--4,8)
Einschränkungen bei Alltagsaktivitäten (ja, stark oder mäßig seit mindestens 6 Monaten)	2,5 (0,7--4,4)	2,4 (-0,1--5,0)	2,8 (0,1--5,5)	1,2 (-3,7--6,1)	4,0 (0,3--7,8)	1,9 (-1,3--5,0)	2,8 (-0,7--6,3)	4,9 (-0,5--10,2)	-4,8 (-10,4--0,8)	3,7 (1,2--6,1)	2,4 (0,1--4,8)
Körperliche Einschränkungen (ja, große Schwierigkeiten oder Unfähigkeit beim Laufen oder Treppen steigen)	-1,2 (-1,8--0,6)	-1,5 (-2,4--0,5)	-0,9 (-1,7--0,0)	0,7 (-0,0--1,5)	-0,4 (-1,4--0,6)	-3,3 (-4,5--2,0)	-2,9 (-4,7--1,1)	4,5 (0,5--8,5)	-2,7 (-4,5--1,0)	-1,2 (-2,0--0,3)	-0,2 (-0,8--0,5)
Diabetes mellitus (ja, jemals ärztlich diagnostiziert)	0,2 (-1,0--1,4)	-0,6 (-1,8--0,6)	0,6 (-1,2--2,4)	0,3 (-1,6--2,2)	0,0 (-1,5--1,6)	-2,1 (-3,9--0,2)	0,3 (-2,2--2,8)	-0,7 (-4,9--3,6)	-3,2 (-5,7--0,6)	-0,2 (-1,6--1,1)	1,2 (0,1--2,3)
Adipositas (BMI ≥ 30 kg/m <sup>2</sup> , berechnet aus Selbstbericht)	1,9 (1,3--2,5)	2,1 (1,3--2,9)	1,7 (0,8--2,5)	2,4 (1,0--3,8)	3,1 (1,8--4,4)	1,0 (0,0--2,0)	2,2 (1,1--3,3)	0,2 (-1,4--1,9)	1,6 (-0,1--3,3)	2,2 (1,4--3,0)	1,0 (0,3--1,8)
Rauchen (ja, täglich oder ja, gelegentlich)	-4,6 (-5,3--3,9)	-3,5 (-4,4--2,6)	-5,7 (-6,7--4,7)	-5,2 (-7,0--3,4)	-4,1 (-5,6--2,7)	-4,7 (-5,8--3,7)	-6,5 (-7,8--5,3)	-1,3 (-2,3--0,3)	-4,1 (-6,0--2,1)	-5,4 (-6,3--4,5)	-2,9 (-3,8--2,0)
Selbsteingeschätzte psychische Gesundheit (Ausgezeichnet/sehr gut)*	-4,6 (-6,3--3,0)	-5,9 (-8,1--3,6)	-3,5 (-5,9--1,1)	-5,2 (-9,9--0,5)	-5,9 (-9,7--2,1)	-4,7 (-7,3--2,1)	-3,2 (-6,0--0,3)	-1,6 (-4,7--1,5)	-0,0 (-5,0--5,0)	-3,6 (-5,9--1,4)	-8,7 (-11,0--6,4)
Depressive Symptomatik (PHQ-8 ≥ 10)	4,2 (2,6--5,9)	6,0 (3,7--8,3)	2,2 (-0,1--4,5)	11,2 (6,4--16,0)	4,4 (0,7--8,0)	2,6 (-0,0--5,2)	2,6 (1,2--4,0)	-0,1 (-5,7--5,6)	3,7 (-1,5--8,9)	3,6 (1,5--5,7)	6,6 (5,5--7,7)
Angstsymptome (GAD-2 ≥ 3)	2,3 (0,5--4,2)	3,2 (0,5--5,9)	1,2 (-1,3--3,6)	9,2 (4,1--14,3)	1,0 (-3,6--5,6)	0,2 (-2,8--3,1)	1,3 (-0,7--3,4)	1,8 (-2,6--6,2)	-1,1 (-7,2--5,1)	3,0 (0,6--5,4)	4,9 (2,6--7,1)
Soziale Unterstützung (starke Unterstützung)	-11,2 (-11,9--10,5)	-12,3 (-13,3--11,3)	-10,0 (-11,0--9,0)	-13,1 (-14,8--11,4)	-14,4 (-15,9--12,8)	-11,2 (-12,3--10,0)	-6,8 (-8,3--5,3)	-2,6 (-5,1--0,1)	-8,6 (-10,5--6,7)	-12,3 (-13,3--11,2)	-14,1 (-15,3--13,0)

Datengrundlage: Alle Surveys beginnend mit GSTel03 bis RKI-Panel 2024, die den jeweiligen Indikator enthalten, inkl. der Methodenstudie GEDA 2024 und der Rekrutierungsstudie des RKI-Panels 2024.

\*Für den Indikator „Selbsteingeschätzte psychische Gesundheit“ wurden aufgrund der im Parallelvergleich identifizierten Fragebogeneffekte bei der Analyse des Trendverlaufs die Angaben aus der Jahreserhebung des RKI-Panels 2024 nicht benutzt.

Analyse mit Alters- und Geschlechtsstandardisierung nach europäischer Standardbevölkerung, für die Stratifizierung nach Bildung auch innerhalb der Bildungsgruppen

Vergleich von Studientyp 1 (EMA-Stichprobe, selbstadministrierter Fragebogen) zu Studientyp 2 (Telefonstichprobe mit Telefoninterview)

Methodeneffekt = modellbasierte Differenz der Prävalenz von Studientyp 1 – Studientyp 2; Methodeneffekt > 0 zeigt an, dass die Prävalenzen in Studientyp 1 höher liegen.

BMI = Body Mass Index, PHQ-8 = Patient Health Questionnaire-8, GAD-2 = Generalized Anxiety Disorder-2



Bei den Indikatoren der körperlichen Gesundheit gibt es kaum Unterschiede nach Geschlecht bei den Methodeneffekten im Zeitverlauf. Die Altersgruppe ab 80 Jahren weist dafür bei fast allen Indikatoren der körperlichen Gesundheit die größten Methodeneffekte auf und diese übersteigen auch die Effekte der psychischen Gesundheit in derselben Altersgruppe. Beim Risikofaktor Adipositas zeigen sich nur geringe Methodeneffekte zwischen den beiden Studientypen über die Zeit und diese unterscheiden sich kaum zwischen den Subgruppen. Beim Rauchen hingegen sind die Effekte stärker ausgeprägt und deutlich größer bei Männern als bei Frauen.

Beim Vergleich der Bildungsgruppen zeigen die Methodeneffekte zunächst kein einheitliches Bild über die Indikatoren. Daher wurde in einer zusätzlichen Analyse auf die bezüglich Bildungsunterschieden aussagekräftigeren Altersgruppen zwischen 30 und 64 Jahren eingeschränkt. Dabei zeigt sich für die Indikatoren der körperlichen und der psychosozialen Gesundheit, dass die Gruppe mit hoher Bildung in den EMA-Stichproben tendenziell häufiger Beschwerden berichtet als in den Telefonstichproben, während die Gruppe mit niedriger Bildung in EMA-Stichproben eher seltener Beschwerden berichtet oder dass geringere Methodeneffekte als in der hohen Bildungsgruppe bestehen. Ausnahmen sind die Risikofaktoren Rauchen und Adipositas, die soziale Unterstützung und die selbsteingeschätzte allgemeine Gesundheit.

#### 4. Diskussion

Mit dem Start des RKI-Panels „Gesundheit in Deutschland“ wurden die bisher vorrangig telefonisch durchgeführten GEDA-Studien des Gesundheitsmonitorings durch ein selbst-administriertes Mixed-Mode-Design (Papier und Online) auf Basis einer Einwohnermeldeamtsstichprobe (EMA-Stichprobe) ersetzt. Der Methodenwechsel soll die Datenqualität verbessern und eine langfristige, kontinuierliche Datenerhebung ermöglichen, sodass Trends zentraler Gesundheitsindikatoren auf Basis von Primärdaten für Deutschland auch zukünftig verlässlich abgebildet werden können. Um die Vergleichbarkeit mit früheren Erhebungen sicherzustellen, wurde eine Parallelerhebung mit identischen Inhalten, aber unterschiedlichen Erhebungsmodi durchgeführt. Es wurden die Stichprobenzusammensetzung, die Prävalenzschätzungen und das Antwortverhalten beider Studien miteinander verglichen und methodenbedingte Unterschiede quantifiziert. Zudem wurde eine Modellierung vorgestellt, mit der die Trends unter Kontrolle des Methodeneffekts geschätzt werden können.

##### Methodeneffekte in der Stichprobenzusammensetzung

Der Vergleich der Parallelerhebungen 2024 deutet auf Auswirkungen des Wechsels vom telefonischen Dual-Frame-Design der GEDA-Studien zu einer selbstadministrierten Mixed-

Mode-Erhebung auf Basis einer Einwohnermeldeamtsstichprobe hin. Das RKI-Panel „Gesundheit in Deutschland“ bildet die Bevölkerung insgesamt verzerrungsärmer ab: junge Erwachsene, Personen mit niedriger Bildung, Einpersonenhaushalte, Hochaltrige und gesundheitlich eingeschränkte Personen sind besser repräsentiert als in den bisherigen telefonischen Studien.

Die Ergebnisse bestätigen den in der Literatur beschriebenen methodischen Vorteil von Einwohnermeldeamtsstichproben gegenüber Telefonstichproben [28,41]. Die vollständigere Abdeckung der Grundgesamtheit und die Möglichkeit einer stratifizierten Zufallsauswahl mit bekannten Auswahlwahrscheinlichkeiten sowie deren Unabhängigkeit von telefonischer Erreichbarkeit oder Ausstattung der Zielperson verringern Selektionsbias und tragen zu einer insgesamt verzerrungsärmeren Stichprobe bei [41]. Vor dem Hintergrund sinkender Teilnahmequoten, die bei Telefonstichproben besonders ausgeprägt sind [7], erscheint die Umstellung auf ein Mixed-Mode-Design folgerichtig und zeitgemäß.

Aktuelle Studien zeigen zudem für Deutschland, dass Mixed-Mode-Erhebungen die geringsten Bildungsverzerrungen aufweisen, während diese in selbstadministrierten Online- und insbesondere in telefonischen Befragungen am stärksten sind [6]. Der Einsatz mehrerer Erhebungsmodi wird daher als zentrale Strategie zur Reduktion systematischer Fehler hervorgehoben [14,15,42]. Durch die parallele Nutzung unterschiedlicher Zugangswege lassen sich Differenzen zwischen Respondern und Nonrespondern verringern und damit der Nonresponse-Bias mildern. Dazu trägt auch die Berücksichtigung unterschiedlicher Präferenzen im Nutzungsverhalten digitaler Technologien bei. Während ältere Personen häufiger Papierfragebögen nutzen [12,43], bevorzugen jüngere und höher gebildete Gruppen digitale Endgeräte [44,45]. Auch die Incentivierung dürfte zur verbesserten Repräsentation beitragen. Die bedingungslose Vergütung in der Rekrutierungsstudie sowie die konditionale Prämie für die Registrierung im RKI-Panel könnten insbesondere für jüngere und einkommensschwächere Personen einen wirksamen Teilnahmeanreiz dargestellt haben.

Eine gezielte Überrekutierung bestimmter Altersgruppen, vor allem von Hochaltrigen ab 80 Jahren, trug zusätzlich zur verbesserten Abbildung dieser Bevölkerungsgruppen bei [23]. Die dort beobachteten höheren Prävalenzen chronischer Erkrankungen und körperlicher Einschränkungen in dieser Altersgruppe sprechen für eine verbesserte Erreichbarkeit gesundheitlich beeinträchtigter Personen, wie auch frühere methodische Studien zeigen [46].

Gleichzeitig können neue Selektionsmechanismen entstehen. Panelerhebungen setzen eine langfristige Teilnahmebereitschaft voraus, die sowohl gesundheitsbewusste als auch gesundheitlich belastete Personen besonders an-

sprechen könnte. Entsprechende Tendenzen wurden auch in anderen Surveys beobachtet, etwa einer Überrepräsentation psychisch belasteter Personen in einem nicht-probabilistischen Online-Panel (d. h. die Teilnehmenden hatten sich aus eigenem Interesse an dem Panel registriert) [14]. Die im RKI-Panel 2024 niedrigere Rauchprävalenz könnte auf eine überproportionale Beteiligung gesundheitsbewussterer Personen hinweisen, auch wenn ein Teil dieses Unterschieds auf Veränderungen in der Gewichtung zurückzuführen ist. Möglich ist, dass bereits bei der Teilnahme an der Rekrutierungsstudie eine Selektion von Gesundheitsbewussteren erfolgte, während die Teilnahme an den telefonischen Studien stärker situativ entschieden wird. Dafür spricht, dass in anderen Studien kein signifikanter Unterschied der Rauchprävalenz zwischen schriftlichen und telefonischen Erhebungen gefunden wurde (z. B. [47]) – teilweise wurde in Telefoninterviews sogar ein günstigeres Gesundheitsverhalten berichtet [12].

Die insgesamt höhere Prävalenz psychischer Belastungen, geringerer sozialer Unterstützung sowie der Anteil chronisch Kranker und körperlich Eingeschränkter bei Hochaltrigen deuten auf eine verbesserte Abbildung gesundheitlich vulnerabler Gruppen hin. Dennoch bleibt unklar, ob diese proportional zu ihrem Bevölkerungsanteil vertreten sind. Erste Befunde deuten auf eine leicht geringere Wiederteilnahme gesundheitlich belasteter Personen hin [23], was auch schon bei der Erstteilnahme am RKI-Panel relevant gewesen sein könnte.

### Methodeneffekte auf Prävalenzschätzer und Antwortverhalten

Im Vergleich der beiden Parallelstudien zeigen sich deutliche Prävalenzunterschiede zwischen den beiden Studiendesigns. Indikatoren der psychischen Gesundheit wie depressive und Angstsymptome oder eine geringe soziale Unterstützung liegen im RKI-Panel deutlich höher als in der GEDA-Telefonerhebung, besonders bei Frauen und bei jüngeren Erwachsenen. Bei Indikatoren der körperlichen Gesundheit sind die Unterschiede geringer ausgeprägt; Adipositas ist moderat häufiger, das Rauchen etwas seltener. Diese Abweichungen sind vor allem im Kontext des Wechsels vom telefonischen Interview zu einer selbstadministrierten Mixed-Mode-Erhebung zu sehen und lassen sich sowohl durch Effekte des Erhebungsmodus als auch durch Unterschiede im Fragebogendesign erklären. Nur beim Rauchen spielt die veränderte Gewichtung eine Rolle.

Abhängig von der Erhebungsmethode können Befragte unterschiedliche Antworten auf dieselbe Frage geben [14, 42, 48], wodurch systematische Verzerrungen, sog. Modeeffekte, entstehen können (z. B. [14, 49]). Auch Sensitivität, Formulierung, Gestaltung und Reihenfolge von Fragen sowie das Format der Antwortoptionen beeinflussen das

Antwortverhalten [27]. Übereinstimmend mit der Literatur (z. B. [12, 39]) zeigen sich die größten Abweichungen bei subjektiven Einschätzungen und Likert-Skalen, wie sie bei der Erfassung von Indikatoren psychischer Gesundheit (z. B. depressiver Symptomatik) verwendet werden. Likert-Skalen sind besonders anfällig für modeabhängige Antwortmuster [26, 39], da kognitive Prozesse und Satisficing (vereinfachte Antwortstrategien ohne Durchdenken der einzelnen Fragen) aufgrund unterschiedlicher Belastung abhängig vom Erhebungsmodus variieren können [28, 48]. Schriftliche Befragungen, die visuell verarbeitet werden, führen häufiger zu Primacy-Effekten (Wahl der zuerst gelesenen Kategorie) [27, 48], zu weniger extremen Antworten [12] sowie zu einer stärkeren Tendenz zur Skalenmitte, also einer neutralen oder mittleren Kategorie [26]. Diese Tendenz zur Mitte ist besonders stark ausgeprägt, wenn die Skala eine klare Mittelkategorie aufweist [26, 27]. Telefonische Interviews, die den Hörsinn ansprechen, begünstigen dagegen Recency-Effekte, also die Wahl der zuletzt gehörten Antwortoption [27, 28, 48, 50]. Je nach Anordnung der Antwortoptionen können sich solche Effekte kumulieren und die Summenwerte verändern. Die für Prävalenzschätzungen notwendige Kategorisierung der Summenwerte kann die Effekte nochmals verstärken [51].

Ein weiterer zentraler Einflussfaktor ist das Ausmaß sozial erwünschten Antwortverhaltens, das mit der wahrgenommenen Privatheit der Befragung variiert. In telefonischen oder persönlichen Interviews können Interviewereffekte auftreten, die zu sozial erwünschten Antworten, Runden (Heaping), extrem positiven Bewertungen und Zustimmungstendenzen (Acquiescence) führen können [14, 19, 26–28, 42, 48]. Befragte neigen in Anwesenheit anderer Personen dazu, sich vorteilhafter darzustellen, Antworten entlang sozialer Normen zu geben und insgesamt positivere Angaben zu machen [9, 26–28, 47–49, 52, 53]. Sozial erwünschte Inhalte werden dabei tendenziell über- und unerwünschte Inhalte hingegen unterberichtet, sodass in interviewbasierten Gesundheitssurveys ein positiver Gesundheitszustand und gesundheitsförderliche Verhaltensweisen überschätzt werden [28]. In selbstadministrierten Befragungen wird dagegen realistischer geantwortet, insbesondere bei sensiblen Themen [9, 14, 28, 52]. Wie bereits für andere psychometrische Inventare oder Einzelfragen festgestellt wurde [12, 48, 54], kann dieses Muster neben kognitiven Verarbeitungsprozessen zur Erklärung herangezogen werden, warum Indikatoren der psychischen Gesundheit im RKI-Panel tendenziell höhere Belastungswerte zeigen. Studien zur Messinvarianz psychometrischer Skalen zeigen, dass unterschiedliche Erhebungsmethoden zwar dieselben Konstrukte messen, sich die Mittelwerte jedoch zwischen den Erhebungsmodi systematisch meist im Zusammenhang mit sozial erwünschtem Antwortverhalten unterscheiden [42, 48]. Die höheren Anteile an Missings bei den beiden Screening-

inventaren von depressiven und Angstsymptomen deuten darauf hin, dass am Telefon mehrfache Likert-Skalen schwerer zu beantworten sind und vermutlich, genauso wie bei den für die Berechnung zur Adipositasprävalenz erforderlichen Gewichtsangaben, auch soziale Erwünschtheit bzw. Scham eine Rolle spielen könnte.

Bei objektiven dichotomen Items, wie der Angabe eines ärztlich diagnostizierten Diabetes, zeigen sich hingegen keine systematischen Effekte des Erhebungsmodus [siehe auch 12, 19, 39]. Im Gegensatz dazu enthalten metrische Angaben wie Körpergewicht häufiger geringe systematische Abweichungen. In schriftlichen Fragebögen wird häufig ein leicht höheres Gewicht angegeben [55], möglicherweise, weil Teilnehmende sich vorab wiegen und weniger sozial erwünscht antworten. Dieser Effekt kann durch Heaping-Effekte wie Auf- und Abrunden im Telefoninterview noch verstärkt werden [56]. Selbst geringe Differenzen können die Adipositasprävalenz in den darauf basierenden BMI-Kategorien um mehrere Prozentpunkte verändern [57] und Gruppenvergleiche beeinflussen, ein bekannter Effekt von Kategorisierungen [51].

Zwischen den selbstadministrierten schriftlichen Modi, also Online- und Papierfragebögen, bestehen laut Metaanalysen kaum Unterschiede im Antwortverhalten [28, 58]. Schriftliche Erhebungen ermöglichen eine gründlichere kognitive Verarbeitung der Fragen, da weniger Zeitdruck besteht und Angaben überprüft oder unterbrochen werden können [15]. Auch die beobachtete Tendenz zur Mitte in schriftlichen Befragungen könnte Ausdruck differenzierterer kognitiver Verarbeitung und stärker reflektierter Beantwortung sein und somit validere Angaben enthalten.

Ein scheinbarer Widerspruch ergibt sich beim Indikator der selbsteingeschätzten psychischen Gesundheit. In der Jahreserhebung des RKI-Panels 2024 berichten mehr Befragte eine sehr gute oder ausgezeichnete psychische Gesundheit, obwohl andere Indikatoren höhere psychische Belastungen zeigen. Der Vergleich mit der Rekrutierungsstudie, in der dieselbe Frage in einem anderen Fragebogenkontext gestellt wurde, spricht für einen Effekt der Positionierung im Fragebogen und für Kontexteffekte (siehe Methodenteil). Dieser Befund verdeutlicht, dass bereits scheinbar kleine methodische Veränderungen, hier der Einsatz der Frage an einer anderen Stelle im Fragebogen, das Antwortverhalten systematisch beeinflussen können [27]. Dies unterstreicht die Bedeutung sorgfältiger Fragebogenkonstruktion sowie gründlicher Testung im Vorfeld, um systematische Fehler (Bias) zu minimieren [49].

Neben klassischen Modeeffekten könnten auch Satisficing-Prozesse eine Rolle spielen, also die Tendenz, kognitiv aufwendige Fragen mit minimalem Aufwand zu beantworten [59]. Dies tritt vor allem bei langen, komplexen oder belastenden Fragebögen auf [26], wie sie im Bereich der psychischen Gesundheit vorkommen. Viele Items zur psy-

chischen Gesundheit ähneln sich inhaltlich und betreffen sensible Themen, was zu Ermüdung oder dem Wunsch nach zügigem Abschluss führen könnte. Erste Analysen zeigen jedoch keine Hinweise auf erhöhte Item-Nonresponse, Abbruchraten oder Straightlining (Ankreuzen immer derselben Antwortkategorie in Fragebatterien). Eine vertiefende Prüfung, auch mit ersten Daten des RKI-Panels 2025, erfolgt derzeit.

Möglich ist zudem, dass die thematische Aufteilung der Fragen auf mehrere Fragebögen die Höhe der Prävalenzen beeinflusst, wenn Personen mit entsprechender Symptomatik gezielter an thematisch passenden Modulen teilnehmen. So wurde berichtet, dass spezifische Erhebungen zum Thema Schmerz zu höheren Schmerzprävalenzen führen als Surveys mit einer umfassenden Gesundheitsthematik [60, 61].

### Unterschiedliche Methodeneffekte in verschiedenen Bevölkerungsgruppen

Die Analysen zeigen, dass sich Methodeneffekte nicht gleichmäßig über alle Bevölkerungsgruppen verteilen. Bei Frauen fallen sie im Zeitverlauf tendenziell stärker aus als bei Männern, mit Ausnahme der Indikatoren zu Einschränkungen bei Alltagsaktivitäten, Diabetes und Rauchen. Besonders ausgeprägt sind die Unterschiede bei jungen Erwachsenen, vor allem bei Indikatoren der psychischen Gesundheit. Diese Altersgruppe ist im RKI-Panel 2024 besser repräsentiert, was sowohl auf eine verbesserte Stichprobenabdeckung als auch auf ein verändertes Antwortverhalten hindeutet. Zugleich zeigt sich bei den Jüngeren die geringste Teilnahmestabilität, wie sinkende Teilnahmeraten über die Teilwellen des RKI-Panels 2024 hinweg belegen [18]. Eine mögliche Selektion innerhalb dieser Altersgruppe, z. B. durch eine stärkere Wiederbeteiligung gesundheitsinteressierter oder besonders belasteter Personen, kann nicht ausgeschlossen werden.

In der Altersgruppe der ab 80-Jährigen treten die größten Methodeneffekte im Bereich der körperlichen Gesundheit auf. Offenbar konnten hier auch stärker eingeschränkte Personen erreicht werden, was sich in höheren Prävalenzen chronischer Erkrankungen und Einschränkungen im RKI-Panel 2024 widerspiegelt. Diese Beobachtung steht im Einklang mit früheren Studien [46, 62], die zeigen, dass zur Einbeziehung hochaltriger und körperlich eingeschränkter Personen aufwendigere Rekrutierungs- und Erhebungsdesigns erforderlich sind, wie sie in der Studie „Gesundheit 65+“ für das bundesweite Monitoring erprobt wurden [63].

Beim Vergleich der Bildungsgruppen zeigten sich bei Einschränkung auf die Altersgruppen zwischen 30 und 64 Jahren, dass in der selbstadministrierten Befragung einer EMA-Stichprobe eine höhere Bildung tendenziell mit einer höheren Prävalenz an Beschwerden einherging, während bei niedrigerer Bildung niedrigere Prävalenzen beobachtet wurden. Ohne Alterseinschränkung überlappen sich diese bildungsspezifischen Methodeneffekte mit Alterseffekten, da der An-

teil von Personen mit niedriger Bildung sowohl bei den 80-Jährigen und Älteren als auch bei jüngeren Personen, die ihre Bildungskarriere noch nicht abgeschlossen haben, besonders hoch ist. Dass dieses Muster sowohl für die körperliche als auch für die psychische Gesundheit zu sehen ist und sowohl für faktische als auch für Einstellungsindikatoren, deutet darauf hin, dass allgemein mehr gesunde Menschen mit niedriger Bildung und bei hoher Bildung mehr Personen mit einer schlechteren Gesundheit in den EMA-Stichproben erreicht wurden. Denkbar wäre auch ein bildungsabhängiges Antwortverhalten, sodass entweder weniger oder mehr Beschwerden abhängig vom Studiendesign berichtet werden. Der Bildungsgradient des Methodeneffektes ist besonders stark bei den Indikatoren der psychosozialen Gesundheit. Eine mögliche Erklärung könnten stärkere Effekte sozialer Erwünschtheit unter Personen mit hoher Bildung sein, wie für Indikatoren von Wohlbefinden gezeigt [64]. Die höher geschätzte Prävalenz schlechter Gesundheit bei Personen mit höherer Bildung und eine verringerte Prävalenz bei Personen mit niedriger Bildung kann dazu führen, dass die im RKI-Panel beobachteten Bildungsunterschiede schwächer ausgeprägt sind als in den vorangegangenen Telefonsurveys.

Die Ergebnisse zeigen, dass methodische Effekte nicht nur das Gesamtniveau der Prävalenzen, sondern auch das Ausmaß gesundheitlicher Ungleichheiten beeinflussen können. Unterschiede in den Prävalenzschätzungen lassen sich nicht eindeutig auf Veränderungen der Stichprobenzusammensetzung oder des Antwortverhaltens zurückführen. Der geschätzte Methodeneffekt umfasst somit alle mit dem Studiendesign verbundenen Veränderungen.

### Schätzung der Methodeneffekte

Durch die Modellierung der Zeitreihen unter Berücksichtigung des Studentyps wurde eine gleichzeitige Schätzung von Trend- und Methodeneffekten ermöglicht. Dabei zeigte sich, dass der Methodeneffekt ein beträchtliches Ausmaß annehmen kann, welches im Einzelfall größer als die Veränderung über die Zeit und generell größer als der zufällige Fehler [65] sein kann. Ohne Berücksichtigung dieser Effekte bestünde die Gefahr, scheinbare Trendbrüche als tatsächliche Gesundheitsveränderungen zu interpretieren.

Zentral für die Modellierung ist die Annahme eines konstanten Methodeneffekts über die Zeit, die sich für die meisten Indikatoren durch gute Anpassung der modellierten Kurven an die gegebenen Datenpunkte als plausibel gezeigt hat (siehe auch Abschnitt 3.4). Dadurch ist es möglich, trotz veränderter Erhebungsmethoden eine inhaltliche Aussage zu Prävalenzentwicklungen über die Zeit zu treffen, auch wenn diese mit größerer Unsicherheit verbunden sind. Bei Indikatoren mit wenigen Datenpunkten für einen der beiden Studentypen oder an den Rändern der Zeitreihe bleibt die Unsicherheit jedoch größer, sodass in diesen Fällen er-

gänzende Sensitivitätsanalysen und eine zurückhaltende Interpretation sinnvoll sind.

Neben der Datenlage ist auch inhaltlich zu prüfen, ob die Annahme eines konstanten Methodeneffekts für den jeweiligen Indikator gerechtfertigt ist. Die Zeitreihen der RKI-Befragungssurveys gehen bis zu 20 Jahre zurück. In diesem Zeitraum kann sich etwa das Ausmaß sozialer Erwünschtheit mit dem Wandel gesellschaftlicher Einstellungen verändert haben. In der schnelllebigen Welt sozialer Medien kann sich sogar innerhalb relativ kurzer Zeit die Akzeptanz von Themen und Einstellungen ändern. Die sinkenden Teilnahmequoten in den Telefonstichproben ([7], [Annex Tabelle 1](#)) können die Stichprobenzusammensetzung und damit die Unterschiede zu EMA-basierten Studien beeinflussen. Die zunehmende Akzeptanz des Online-Fragebogens etwa auch in den höheren Altersgruppen kann die Anteile der einzelnen Erhebungsmethoden im Mixed-Mode-Design über die Zeit verschieben und damit Prävalenzschätzungen verändern. Die Annahme eines konstanten Methodeneffekts stellt damit immer eine gewisse Vereinfachung dar. Nichtsdestotrotz kann sie hilfreich sein, wenn sie durch die Datenlage gestützt wird, um die Größenordnung von Methodeneffekten abzuschätzen.

### Stärken und Limitationen

Methodeneffekte sind kein singuläres Phänomen der RKI-Gesundheitssurveys, sondern betreffen alle auf freiwilligen Befragungen beruhenden Studien. Der Neustart des RKI-Gesundheitspanels ermöglichte es jedoch, durch die parallel durchgeführte Erhebung mit einem Design analog zu den vorangegangenen Studien bzw. die Modellierung über die gesamte Zeitreihe aus zwei verschiedenen Studiendesigns, den Methodeneffekt zu schätzen und gleichzeitig die Entwicklung der interessierenden Prävalenzen über die Zeit zu beurteilen. Eine zentrale Voraussetzung dafür war, dass bereits in einer früheren Erhebung 2014/2015 ein mit dem RKI-Panel vergleichbares Design eingesetzt worden war. Da es sich hier jedoch nur um einen einzigen früheren Datenpunkt mit diesem Design handelt, sind die geschätzten Methodeneffekte als Annäherung an die Auswirkungen des Methodenwechsels zu verstehen, unter der Annahme eines über die Zeit konstanten Methodeneffekts.

Die Analysen deuten darauf hin, dass Einwohnermeldeamtsstichproben mit selbstadministrierter Befragung wie das RKI-Panel 2024 die Bevölkerungsstruktur realistischer abbilden als Telefonsurveys. Dennoch bleiben Verzerrungsquellen bestehen, vor allem eine geringere Teilnahmebereitschaft von Personen mit niedriger Bildung. In der Gewichtung konnten jedoch nur die grundlegenden soziodemografischen Merkmale Alter, Geschlecht, Region, Gemeindetyp, Bildung und Haushaltsgröße berücksichtigt werden. Darüber hinaus erfolgte kein systematischer Abgleich mit weiteren Referenz-



daten. Somit können potenzielle Verzerrungen, wie sie allgemein in Querschnittstudien auftreten, etwa in Bezug auf politische Einstellungen, Gesundheitsbewusstsein oder Teilnahmemotivation, nicht ausgeschlossen werden.

Wie in den Telefonstudien wurden auch im RKI-Panel 2024 nur deutschsprachige Erhebungsinstrumente eingesetzt. Dies führt dazu, dass Menschen, deren Präferenzsprache nicht Deutsch ist, nicht ideal abgebildet werden können. Um dieser Verzerrung perspektivisch entgegen zu wirken, prüfen wir derzeit den Einsatz ausgewählter übersetzter Materialien und Fragebögen für künftige Erhebungswellen und insbesondere Rekrutierungsstudien.

Da die parallel erhobenen Studien nicht auf identischen Individuen beruhen, lassen sich Veränderungen im Antwortverhalten nur durch den Vergleich der Häufigkeit von Antwortkategorien, nicht aber nicht intra-individuell beurteilen. Zudem standen für einige Indikatoren bislang nur wenige Messzeitpunkte zur Verfügung, was die Robustheit der Modellierungen einschränkt. Insgesamt gab es selbst bei längeren Zeitreihen nur wenige Datenpunkte, die auf EMA-Stichproben mit schriftlicher Befragung beruhten. Besonders an den Rändern des Beobachtungszeitraums können Trends nur bedingt approximiert werden. Zukünftige Erhebungen werden daher entscheidend sein, um beobachtete Trendänderungen zu bestätigen oder zu widerlegen.

Das vorgestellte Verfahren ermöglicht es, den Methodeneffekt für jeden Indikator zu quantifizieren und damit Prävalenzschätzungen aus unterschiedlichen Studiendesigns vergleichbar zu machen. Dies schafft eine Grundlage, um Prävalenzen aus verschiedenen Studiendesigns einzuordnen, auch über die RKI-Studien hinaus.

## Schlussfolgerungen

Die Etablierung einer Panel-Infrastruktur mit überwiegend kostengünstiger Onlinebefragung kann für das Gesundheitsmonitoring zukünftig eine methodisch konstante Datenerhebung ermöglichen, die eine Grundvoraussetzung für valide Trendaussagen ist [11–13]. Die Stichprobe im RKI-Panel 2024 bildet die Bevölkerungsstruktur hinsichtlich Alter, Geschlecht, Haushaltsgröße und Bildungsstand differenzierter ab als die früheren GEDA-Studien. Erste Hinweise deuten darauf hin, dass Personen mit schlechterem Gesundheitszustand oder erhöhter psychischer Belastung eher teilnehmen als in telefonischen Erhebungen, in denen sie tendenziell unterrepräsentiert sind. Beim Rauchverhalten deutet sich dagegen eine leicht erhöhte Teilnahme von gesundheitsbewussten Personen an, was zu Verzerrungen führen kann. Wie sich diese Effekte insgesamt auswirken, lässt sich derzeit nicht abschließend beurteilen.

Das Panel-Design bringt im Vergleich zu Querschnitterhebungen die Herausforderung der Panelattrition mit sich, also dem systematischen Ausstieg von Teilnehmenden im

Zeitverlauf. Dieser Entwicklung kann durch gezielte Teilnahmeanreize, Panelpflege und Gewichtungskorrekturen sowie Auffrischungsstichproben entgegengewirkt werden, dennoch bleibt sie angesichts sinkender Teilnahmequoten in Bevölkerungsbefragungen (z. B. [8]) eine zentrale Herausforderung für die Datenqualität und Repräsentativität.

Wie bereits in der Methodenstudie für den Methodenwechsel auf GEDA 2014/2015 gesehen [19], kann das selbst-administrierte Mixed-Mode-Design (Papier und Online) im RKI-Panel durch den Wegfall von Interviewereffekten sowie die größere Privatheit insbesondere bei sensiblen Fragen zu validen Daten führen [14, 58]. Schriftliche Befragungen bieten zudem zeitliche Flexibilität, Antworten können unterbrochen, überdacht oder recherchiert werden. Gleichzeitig steigen damit die Anforderungen an Fragebogengestaltung und Dramaturgie. Da dieselben Personen wiederholt befragt werden, sind eine zumutbare Befragungslänge [66], klare Struktur und hohe Verständlichkeit entscheidend, um Satisficing zu vermeiden. Ein sorgfältiger Pretest, regelmäßige Zufriedenheitsbefragungen sowie angemessene Teilnahmeanreize sind daher wichtige Bestandteile der Qualitätssicherung.

Die Analysen verdeutlichen die zentrale Bedeutung methodischer Konstanz über die Zeit, sowohl hinsichtlich des Studiendesigns (Stichprobenziehung, Erhebungsmethode) als auch der Messinstrumente (Frage- und Antwortformulierung, Position im Fragebogen). Wo Änderungen unvermeidbar sind, sollten sie durch Methodenstudien begleitet werden, um mögliche Effekte zu quantifizieren. Dies gilt insbesondere für psychische Gesundheitsindikatoren, die besonders anfällig für Modus- und Kontexteffekte sind. Nur so können beobachtete Unterschiede als inhaltliche Veränderungen und nicht als Messartefakte interpretiert werden. Die hier eingesetzten statistischen Verfahren ermöglichen es, Methodeneffekte zu identifizieren und zu schätzen.

## Datenschutz und Ethik

Durch den Bundesbeauftragten für Datenschutz (BDS) wurde für die Studien GSTel03 und GSTel06 Zustimmung erteilt. Die GEDA-Studien sowie das Panel „Gesundheit in Deutschland“ 2024 unterliegen der strikten Einhaltung der datenschutzrechtlichen Bestimmungen der EU-Datenschutzgrundverordnung (DSGVO) und des Bundesdatenschutzgesetzes (BDSG). Die Ethikkommission der Charité – Universitätsmedizin Berlin hat die Studie GEDA 2019/2020-EHIS (Antragsnummer EA2/070/19) sowie die Studien GEDA 2022/2023 (Antragsnummer EA2/201/21) unter ethischen Gesichtspunkten geprüft und der Durchführung der Studienvorhaben zugestimmt. Die Ethik-Kommission der Ärztekammer Berlin hat die Studie „Gesundheit in Deutschland“ (RKI-Panel) unter ethischen Gesichtspunkten gemäß der aktuellen Fassung der Deklaration von Helsinki geprüft und der Durchführung des Studienvorhabens und der Jahreswelle 2024 zugestimmt (Antragsnummer Eth-63/23, Amendement 1). Die Teilnahme an den Studien war freiwillig. Die Teilnehmenden wurden über die Ziele und Inhalte der Studie sowie über den Datenschutz informiert und gaben ihre mündliche bzw. schriftliche (entweder elektronisch im Rahmen der Online-Befragung oder schriftlich) Einwilligung (informed consent).

**Datenverfügbarkeit**

Die Autorinnen und Autoren geben an, dass für die den Ergebnissen zugrunde liegenden Daten Zugangsbeschränkungen bestehen. Die Surveys GEDA 2009, GEDA 2010, GEDA 2012, GEDA 2014/2015-EHIS und GEDA 2019/2020-EHIS können über das Forschungsdatenzentrum (FDZ) des Robert Koch-Instituts als Scientific Use Files beantragt werden. Die anderen verwendeten Surveys sind im Rahmen einer erweiterten Datennutzung über das FDZ bei begründetem Forschungsinteresse zugänglich. Anfragen können per E-Mail an [fdz@rki.de](mailto:fdz@rki.de) gestellt werden. Weitere Informationen finden sich auf der folgenden Webseite: [www.rki.de/suf](http://www.rki.de/suf)

**Förderungsinweis**

Das Panel „Gesundheit in Deutschland“ wurde mit Mitteln des Robert Koch-Instituts und des Bundesministeriums für Gesundheit finanziert.

**Beiträge der Autorinnen und Autoren**

Maßgebliche Beiträge zu Konzeption oder Design der Arbeit: EM, ASR, FV; zur Erhebung der verwendeten Daten: JL, JA, SD, RK; zur Analyse der verwendeten Daten: ASR, FV; zur Interpretation der verwendeten Daten: FV, ASR, NM, SD, EM, AS, BG, JB, YD, RS, RK. Ausarbeitung des Manuskripts: EM, FV, ASR; kritische Überarbeitung bedeutender Inhalte: EM, FV, ASR, AS, BG, NM, SD, JB, YD, RS, JA, JL, RK. Finale Version des Manuskripts gelesen und der Veröffentlichung zugestimmt: EM, FV, ASR, NM, SD, AS, BG, JB, YD, RS, JA, JL, RK.

**Interessenkonflikt**

Die Autorinnen und Autoren geben an, dass kein Interessenkonflikt besteht.

**Danksagung**

Wir danken Petra Rattay für die wertvollen Hinweise und Diskussionsbeiträge zum Indikator soziale Unterstützung, Sabine Born und ihrem Team für die umfassende und sorgfältige Datenaufbereitung, Stephan Junker für die Hinweise zu methodischen Fragestellungen sowie Carmen Koschollek für die Klärung von migrationsbezogenen Fragen. Weiterhin danken wir den Reviewern für ihre Anregungen, die das Manuskript klarer strukturiert und bereichert haben. ChatGPT (GPT-5, OpenAI) wurde für Unterstützung bei der sprachlichen und stilistischen Überarbeitung einzelner Textpassagen genutzt, Consensus (<https://consensus.app/>) in der kostenlosen Version für Unterstützung bei der Suche nach geeigneten Referenzen. Die Auswahl der Referenzen sowie alle Entscheidungen über Inhalte und Formulierungen wurden von den Autorinnen und Autoren getroffen.

**Literatur**



















- Kurth BM, Lange C, Kamtsiuris P, Hölling H. Gesundheitsmonitoring am Robert Koch-Institut. Bundesgesundheitsbl. 2009;52(5):557–70. doi: 10.1007/s00103-009-0843-3.
- Webportal der Gesundheitsberichterstattung des Bundes. Robert Koch-Institut; [cited 07.11.2025]. Available from: <https://www.gbe.rki.de/>.
- Lemcke J, Loss J, Allen J, Öztürk I, Hintze M, Damerow S, et al. Gesundheit in Deutschland: Etablierung eines bevölkerungsbezogenen Gesundheitspanels. J Health Monit. 2024(S2):2–22. doi: 10.25646/11913.2.
- Christensen AI, Lau CJ, Kristensen PL, Johnsen SB, Wingstrand A, Friis K, et al. The Danish National Health Survey: Study design, response rate and respondent characteristics in 2010, 2013 and 2017. Scand J Public Health. 2022;50(2):180–8. doi: 10.1177/1403494820966534.
- Wagner M, Kuppler M, Rietz C, Kaspar R. Non-response in surveys of very old people. Eur J Ageing. 2019;16(2):249–58. Epub 20180903. doi: 10.1007/s10433-018-0488-x.
- Stein A, Gummer T, Naumann E, Rohr B, Silber H, Auriga R, et al. Education bias in probability-based surveys in Germany: evidence and possible solutions. Int J Soc Res Methodol. 2025;1–18. doi: 10.1080/13645579.2025.2508889.
- Luiten A, Hox J, de Leeuw E. Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys. J Off Stat. 2020;36(3):469–87. doi: 10.2478/jos-2020-0025.
- Jabkowski P, Cichocki P. Survey response rates in European comparative surveys: a 20-year decline irrespective of sampling frames or survey modes. Quality & Quantity. 2025;59(1):635–55. doi: 10.1007/s11135-024-01993-9.
- Olson K, Smyth JD, Horwitz R, Keeter S, Lesser V, Marken S, et al. Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys: AAPOR Task Force Report. J Surv Stat Methodol. 2020;9(3):381–411. doi: 10.1093/jssam/srmz062.
- Lange C, Finger JD, Allen J, Born S, Hoebel J, Kuhnert R, et al. Implementation of the European health interview survey (EHIS) into the German health update (GEDA). Arch Public Health. 2017;75(1):40. doi: 10.1186/s13690-017-0208-6.
- Zinn S, Wolbring T. Recent Methodological Advances in Panel Data Collection, Analysis, and Application. Surv Res Methods. 2023;17(3):219–22. doi: 10.18148/srm/2023.v17i3.8317.
- Christensen AI, Ekholm O, Glümer C, Juel K. Effect of survey mode on response patterns: comparison of face-to-face and self-administered modes in health surveys. Eur J Public Health. 2013;24(2):327–32. doi: 10.1093/eurpub/ckt067.
- Firebaugh G. Analyzing Repeated Surveys. Thousand Oaks, California: SAGE Publications, Inc. 1997 [cited 14.08.2025]. Available from: <https://methods.sagepub.com/book/mono/analyzing-repeated-surveys/toc>.
- DeLeeuw ED. Mixed-Mode: Past, Present, and Future. Survey Research Methods. 2018;12(2):75–89. doi: 10.18148/srm/2018.v12i2.7402.
- Dillman DA, Smyth JD, Christian LM. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th Edition. New York: Wiley 2014.
- Lemcke J, Öztürk I, Damerow S, Heller T, Born S, Wetzstein M, et al. Recruitment of a probability-based general population health panel for public health research in Germany: The panel ‘Health in Germany’. BMC Med Res Methodol. Under review.
- AAPOR. The American Association for Public Opinion Research. Final Dispositions of Case Codes and Outcome Rates for Surveys. 10th edition. 2023 Standard Definitions [cited 07.11.2025]. Available from: <https://aapor.org/wp-content/uploads/2024/03/Standards-Definitions-10th-edition.pdf>.
- Lemcke J, Damerow S, Kuttig T, Öztürk I, Albrecht S, Heller T, et al. Panel „Gesundheit in Deutschland“ des Robert Koch-Instituts – Hintergrund und Methode der Jahrerhebung 2024. J Health Monit. 2025;10(4):e13546. doi: 10.25646/13546.
- Hoebel J, von der Lippe E, Lange C, Ziese T. Mode differences in a mixed-mode health interview survey among adults. Arch Public Health. 2014. doi: 10.1186/2049-3258-72-46.
- Mauz E, von der Lippe E, Allen J, Schilling R, Müters S, Hoebel J, et al. Mixing modes in a population-based interview survey: comparison of a sequential and a concurrent mixed-mode design for public health research. Arch Public Health. 2018;76(1):8. doi: 10.1186/s13690-017-0237-1.
- Behrens K, Böltken F, Dittmar H, Götsche F, Gutfleisch R, Habla H, et al. Regionale Standards: Ausgabe 2019. GESIS-Schriftreihe, 23: GESIS – Leibniz-Institut für Sozialwissenschaften doi: 10.12141/ssor.62343.
- Brauns H, Scherer S, Steinmann S. The CASMIN educational classification in international comparative research. In: Hoffmeyer-Zlotnik JHP, Wolf C, editors. Advances in cross-national comparison. Boston, MA: Springer; 2003. p. 221–44.
- Damerow S, Kuhnert R, Schaffrath Rosario A, Lemcke J. Weighting Strategy and Selection Analysis in the Panel ‘Health in Germany’: Methods and Results for the 2024 Annual Survey. BMC Med Res Methodol. Forthcoming.

- 24 Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. (ADM), Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e. V. (ASI), Statistisches Bundesamt. Demografische Standards. GESIS – Leibniz-Institut für Sozialwissenschaften. 2024 [cited 11.06.2025]. Available from: [https://www.destatis.de/DE/Methoden/Demografische-Regionale-Standards/Downloads/demografische-standards-2024.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Demografische-Regionale-Standards/Downloads/demografische-standards-2024.pdf?__blob=publicationFile).
- 25 UNESCO Institute for Statistics. International Standard Classification of Education: ISCED 2011. Montreal, Canada: UIS; 2012.
- 26 d'Ardenne J, Collins D, Gray M, Jessop C, Pilley S. Assessing the risk of mode effects. Review of proposed survey questions for waves 7-10 of Understanding Society. Working Paper Series [Internet]. 2017 [cited 19.08.2025]. Available from: <https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2017-04.pdf>.
- 27 Bogner K, Landrock U. Antworttendenzen in standardisierten Umfragen. GESIS Survey Guidelines. Mannheim: GESIS Leibniz Institut für Sozialwissenschaften; 2015.
- 28 Bowling A. Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*. 2005;27(3):281–91. doi: 10.1093/pubmed/fdi031.
- 29 Cox B, van Oyen H, Cambois E, Jagger C, le Roy S, Robine JM, et al. The reliability of the Minimum European Health Module. *Int J Public Health*. 2009;54(2):55–60. doi: 10.1007/s00038-009-7104-y.
- 30 Robine JM, Jagger C. Creating a coherent set of indicators to monitor health across Europe: the Euro-REVES 2 project. *Eur J Public Health*. 2003;13(3 Suppl):6–14. doi: 10.1093/eurpub/13.suppl\_1.6.
- 31 Eurostat. European Health Interview Survey (EHIS wave 3) – Methodological manual. Luxembourg: Publications Office of the European Union. 2018 [cited 09.07.2025]. Available from: <https://ec.europa.eu/eurostat/documents/3859598/8762193/KS-02-18-240-EN-N.pdf>.
- 32 World Health Organization (WHO). Obesity: preventing and managing the global epidemic: report of a WHO consultation. Geneva: WHO; 2000 [cited 08.07.2025]. Available from: [https://cfmws.ca/CFMWS/media/Kingston/WHO\\_TRS\\_894.pdf&ved=2ahUKewiH-sazc6IORAxUjSvEDHZydkgcQFnoECCMQAQ&usq=AOvVaw1cl-0H08d\\_JqG-M9DRIm4V1](https://cfmws.ca/CFMWS/media/Kingston/WHO_TRS_894.pdf&ved=2ahUKewiH-sazc6IORAxUjSvEDHZydkgcQFnoECCMQAQ&usq=AOvVaw1cl-0H08d_JqG-M9DRIm4V1).
- 33 Starker A, Schienkiewitz A, Damerow S, Kuhnert R. Verbreitung von Adipositas und Rauchen bei Erwachsenen in Deutschland – Entwicklung von 2003 bis 2023. *J Health Monit*. 2025;10(1):1–14. doi: 10.25646/12990.
- 34 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. Epub 2014/09/19. doi: 10.1186/1472-6963-14-398.
- 35 Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord*. 2009;114(1–3):163–73. doi: 10.1016/j.jad.2008.06.026.
- 36 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(5):317–25. doi: 10.7326/0003-4819-146-5-200703060-00004.
- 37 Kocalevent RD, Berg L, Beutel ME, Hinz A, Zenger M, Härter M, et al. Social support in the general population: standardization of the Oslo social support scale (OSSS-3). *BMC Psychology*. 2018;6(1):31. doi: 10.1186/s40359-018-0249-9.
- 38 Statistisches Amt der Europäischen Union (Eurostat). Revision of the European Standard Population. Report of Eurostat's task force. Luxembourg: Publications Office of the European Union; 2013 [cited 07.06.2024]. Available from: <https://ec.europa.eu/eurostat/de/web/products-manuals-and-guidelines/-/KS-RA-13-028>.
- 39 Maslovskaya O, Calderwood L, Ploubidis G, Nicolaas G. Adjustments for Mode Effects. Southampton 2023.
- 40 Schwarz G. Estimating the Dimension of a Model. *Annals of Statistics*. 1978;6:461–4. doi: 10.1214/aos/1176344136.
- 41 Häder S. Sampling in Practice. GESIS Survey Guidelines. Mannheim, Germany; 2016. doi: 10.15465/gesis-sg\_en\_014.
- 42 Zager Kocjan G, Lavtar D, Sočan G. The effects of survey mode on self-reported psychological functioning: Measurement invariance and latent mean comparison across face-to-face and web modes. *Behavior Research Methods*. 2023;55(3):1226–43. doi: 10.3758/s13428-022-01867-8.
- 43 Gaertner B, Scheidt-Nave C, Koschollek C, Fuchs J. Gesundheitliche Lage älterer und hochaltriger Menschen in Deutschland: Ergebnisse der Studie Gesundheit 65+. *J Health Monit*. 2023(3):7–31. doi: 10.25646/11564.
- 44 Haan M, Lugtig P, Toepoel V. Can we predict device use? An investigation into mobile device use in surveys. *Int J Soc Res Methodol*. 2019;22(5):517–31. doi: 10.1080/13645579.2019.1593340.
- 45 Gummer T, Höhne JK, Rettig T, Roßmann J, Kummerow M. Is there a growing use of mobile devices in web surveys? Evidence from 128 web surveys in Germany. *Quality & Quantity*. 2023;57(6):5333–53. doi: 10.1007/s11135-022-01601-8.
- 46 Gaertner B, Lüdtke D, Koschollek C, Grube M, Baumert J, Scheidt-Nave C, et al. Effects of a sequential mixed-mode design on participation, contact and sample composition – Results of the pilot study “IMOA – Improving Health Monitoring in Old Age”. *Survey Methods: Insights from the Field (SMIF)*. 2019.
- 47 Feveile H, Olsen O, Høgh A. A randomized trial of mailed questionnaires versus telephone interviews: Response patterns in a survey. *BMC Med Res Methodol*. 2007;7(1):27. doi: 10.1186/1471-2288-7-27.
- 48 Sakshaug J, Cernat A, Silverwood RJ, Calderwood L, Ploubidis GB. Measurement Equivalence in Sequential Mixed-Mode Surveys. *Survey Research Methods*. 2022;16(1):29–43. doi: 10.18148/srm/2022.v16i1.7811.
- 49 Schouten B, van den Brakel J, Buelens B, van der Laan J, Klausch T. Disentangling mode-specific selection and measurement bias in social surveys. *Social Science Research*. 2013;42(6):1555–70.
- 50 Krosnick JA, Alwin DF. An evaluation of a cognitive theory of response-order effects in survey measurement. *Am J Epidemiol*. 1987;125(2):201–19.
- 51 Flegal KM, Keyl PM, Nieto FJ. Differential Misclassification Arising from Nondifferential Errors in Exposure Measurement. *Am J Epidemiol*. 1991;134(10):1233–46. doi: 10.1093/oxfordjournals.aje.a116026.
- 52 Kreuter F, Presser S, Tourangeau R. Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly*. 2008;72(5):847–65. doi: 10.1093/poq/nfn063.
- 53 Ye C, Fulton J, Tourangeau R. More positive or More Extreme? A Meta-Analysis of Mode Differences in Response Choice. *Public Opinion Quarterly*. 2011;75(2):349–65. doi: 10.1093/poq/nfn009.
- 54 Zhang X, Kuchinke L, Woud ML, Velten J, Margraf J. Survey method matters: Online/offline questionnaires and face-to-face or telephone interviews differ. *Computers in Human Behavior*. 2017;71:172–80. doi: 10.1016/j.chb.2017.02.006.
- 55 Lin CJ, DeRoo LA, Jacobs SR, Sandler DP. Accuracy and reliability of self-reported weight and height in the Sister Study. *Public Health Nutr*. 2012;15(6):989–99. Epub 20111209. doi: 10.1017/S1368980011003193.
- 56 García-González JM, Martín-Criado E. Rounding as an indicator of bias in reported body weight in health surveys. *J Biosoc Sci*. 2021;53(5):790–9. Epub 2020/09/15. doi: 10.1017/S002193202000053X.
- 57 Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev*. 2007;8(4):307–26. doi: 10.1111/j.1467-789X.2007.00347.x.
- 58 Dodou D, de Winter JCF. Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior*. 2014;36:487–95. doi: 10.1016/j.chb.2014.04.005.
- 59 Krosnick JA, Narayan S, Smith WR. Satisficing in surveys: Initial evidence. *New Directions for Evaluation*. 1996;1996(70):29–44. doi: 10.1002/ev.1033.

- 60 Papageorgiou AC, Croft PR, Ferry S, Jayson MI, Silman AJ. Estimating the prevalence of low back pain in the general population. Evidence from the South Manchester Back Pain Survey. *Spine (Phila Pa 1976)*. 1995;20(17):1889–94. doi: 10.1097/00007632-199509000-00009.
- 61 Neuhauser H, Ellert U, Ziese T. Chronische Rückenschmerzen in der Allgemeinbevölkerung in Deutschland 2002/2003: Prävalenz und besonders betroffene Bevölkerungsgruppen. *Gesundheitswesen*. 2005;67(10):685–93. doi: 10.1055/s-2005-858701.
- 62 Gaertner B, Fuchs J, Scheidt-Nave C. Gesundheit im Alter und deren Bedeutung für Teilhabe: Expertise zum Neunten Altersbericht der Bundesregierung 2025. Berlin: Deutsches Zentrum für Altersfragen. 2025 [cited 10.07.2025]. Available from: <https://www.neunter-altersbericht.de/expertisen>.
- 63 Fuchs J, Gaertner B, Perlitz H, Kuttig T, Klingner A, Baumert J, et al. Studie zur Gesundheit älterer Menschen in Deutschland (Gesundheit 65+): Zielsetzung, Konzeption und Durchführung. *J Health Monit*. 2023;8(3):66–90. doi: 10.25646/11662.
- 64 Piccitto G, Liefbroer AC, Emery T. Does the Survey Mode Affect the Association Between Subjective Well-being and its Determinants? An Experimental Comparison Between Face-to-Face and Web Mode. *J Happiness Stud*. 2022;23(7):3441–61. doi: 10.1007/s10902-022-00553-y.
- 65 Greenland S, Pearce N. Statistical foundations for model-based adjustments. *Annu Rev Public Health*. 2015;36:89–108. doi: 10.1146/annurev-publhealth-031914-122559.
- 66 Galesic M, Bosnjak M. Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*. 2009;73(2):349–60. doi: 10.1093/poq/nfp031.



**Annex Tabelle 1:** Übersicht über die Methodik der genutzten Befragungssurveys

Survey	Stichprobe		Erhebungsmodus	Feldzeit	Fallzahl <sup>a</sup>	Ausschöpfungsquote <sup>b</sup>	Gewichtung			
	Einwohnermeldeamts (EMA)-Stichprobe	Telefonstichprobe: Festnetz (F)/Dual-Frame (DF)	Telefoninterview  / Mixed-Mode schriftlich (webbasiert  /Papier  )		Anzahl Teilnehmende ab 18 Jahren		Bevölkerungsstruktur Stichtag 31.12.	Bildung/Anteil Single-HH laut Mikrozensus	Anpassungsgewichtung basiert auf ...	Bildung klassifiziert nach ...
GSTel03 (2003) <a href="#">[1,2]</a>		F		9/2002 bis 3/2003	8.318	In Anlehnung RR2: 53,8 %	2002	2003	Geschlecht, Alter, Region, Bildung	ISCED 1997 <a href="#">[18]</a>
GSTel06 (2006) <a href="#">[3]</a>		F		10/2005 bis 3/2006	5.542	In Anlehnung RR2: 37,4 %	2006	2007	Geschlecht, Alter, Region, Bildung	ISCED 1997 <a href="#">[18]</a>
GEDA 2009 <a href="#">[4–6]</a>		F		7/2008 bis 5/2009	21.262	RR2: 25,7 %	2007	2007	Geschlecht, Alter, Bundesland, Bildung	ISCED 1997 <a href="#">[18]</a>
GEDA 2010 <a href="#">[4,6,7]</a>		F		9/2009 bis 7/2010	22.050	RR2: 25,3 %	2008	2008	Geschlecht, Alter, Bundesland, Bildung	ISCED 1997 <a href="#">[18]</a>
GEDA 2012 <a href="#">[4,6,8]</a>		F		3/2012 bis 2/2013	19.294	RR2: 20,5 %	2011	2010	Geschlecht, Alter, Bundesland, Bildung	ISCED 1997 <a href="#">[18]</a>
GEDA 2014/2015-EHIS <a href="#">[4,9,10]</a>	EMA		 	11/2014 bis 7/2015	24.016	RR1: 27,6 %	2014	2013	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (Kreistyp)	ISCED 2011 <a href="#">[19]</a>
GEDA 2019/2020-EHIS <a href="#">[4,11]</a>		DF		4/2019 bis 9/2020	22.708	RR3: 21,6 %	2019	2017	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (Kreistyp)	ISCED 2011 <a href="#">[19]</a>
GEDA 2022 <a href="#">[4,12]</a>		DF		2/2022 bis 1/2023	33.149 <sup>c</sup>	RR3: 16,1 % – 19,6 %	2020	2018	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (Kreistyp) [Gewichtung pro Erhebungszeitpunkt und ggf. Modul]	ISCED 2011 <a href="#">[19]</a>
GEDA 2023 <a href="#">[4,12]</a>		DF		1/2023 bis 2/2024	30.002 <sup>c</sup>	RR3: 17,4 % – 19,8 %	2020	2018	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (Kreistyp) [Gewichtung pro Erhebungszeitpunkt und ggf. Modul]	ISCED 2011 <a href="#">[19]</a>
RKI-Panel 2024 Rekrutierungsstudie <a href="#">[13–15]</a>	EMA		 	1/2024 bis 5/2024	61.460	RR2: 37,6 %	2023	2021	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (BIK-Gemeindetyp <a href="#">[22]</a> ), Anteil Single-Haushalte <a href="#">[16]</a>	CASMIN <a href="#">[20,21]</a>
GEDA 2024 (Methodenstudie, Parallelerhebung) <a href="#">[4]</a>		DF		6/2024 bis 9/2024	4.016	RR3: 19,6 % – 20,0 %	2023	2021	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (Kreistyp)	ISCED 2011 <a href="#">[19]</a>
RKI-Panel 2024 <a href="#">[13,14,17]</a>	EMA		 	6/2024 bis 1/2025	FB A: 27.199 FB B: 27.147 FB C: 27.102 FB D: 27.306	RR2: 77,7 % – 78,1 % RR2cum: 22,3 % – 22,4 %	2023	2021	Geschlecht, Alter, Bundesland, Bildung, Siedlungsstruktur (BIK-Gemeindetyp <a href="#">[22]</a> ), Anteil Single-Haushalte; zusätzlich Ausgleich von Drop-out in den Teilwellen anhand der Angaben aus der Rekrutierungsstudie <a href="#">[16]</a>	CASMIN <a href="#">[20,21]</a>

FB=Fragebogen. Der Begriff „Region“ steht jeweils für die Zusammenfassung von Bundesländern; diese ist in den einzelnen Erhebungen unterschiedlich erfolgt.

<sup>a</sup>Die hier angegebenen Fallzahlen beziehen sich auf die Altersgruppe ab 18 Jahren.

<sup>b</sup>Ausschöpfungsquoten RR1, RR2, RR3 sowie RR2cum nach AAPOR [\[23\]](#). Im RKI-Panel steht RR2 für die Wiederteilnahmequote der registrierten Panelisten und RR2cum für die kumulative Ausschöpfungsquote bezogen auf die zur Rekrutierungsstudie Eingeladenen.

<sup>c</sup>GEDA 2022 und GEDA 2023: Gesamtzahl der Teilnehmenden über alle Erhebungszeitpunkte und alle Module

## Annex Tabelle 1: Literatur

- 1 Robert Koch-Institut. Erster telefonischer Gesundheitssurvey des Robert Koch-Instituts – Methodische Beiträge. Berlin: RKI; 2005.
- 2 Robert Koch-Institut. Telefonischer Gesundheitssurvey (2003) – GSTel03, Studien-ID REL1\_119 Köln: Data Reference (GESIS); 2005.
- 3 Robert Koch-Institut. Telefonischer Gesundheitssurvey (2006) – GSTel06, Studien-ID REL1\_122. Köln: Data Reference (GESIS); 2006.
- 4 Robert Koch-Institut. GEDA: Gesundheit in Deutschland aktuell. 2024 [cited 10.11.2025]. Available from: [https://www.rki.de/DE/Themen/Nichtuebertragbare-Krankheiten/Studien-und-Surveillance/Studien/GEDA/geda\\_gesamt.html](https://www.rki.de/DE/Themen/Nichtuebertragbare-Krankheiten/Studien-und-Surveillance/Studien/GEDA/geda_gesamt.html).
- 5 Robert Koch-Institut. Daten und Fakten: Ergebnisse der Studie „Gesundheit in Deutschland aktuell 2009“. Berlin: RKI; 2011.
- 6 Lange C, Jentsch F, Allen J, Hoebel J, Kratz AL, von der Lippe E, et al. Data Resource Profile: German Health Update (GEDA) – the health interview survey for adults in Germany. *Intern J Epidemiol* 2015;44(2): 442–450. doi: 10.1093/ije/dyv067.
- 7 Robert Koch-Institut. Daten und Fakten: Ergebnisse der Studie „Gesundheit in Deutschland aktuell 2010“. Berlin: RKI; 2012.
- 8 Robert Koch-Institut. Daten und Fakten: Ergebnisse der Studie „Gesundheit in Deutschland aktuell 2012“. Berlin: RKI; 2014.
- 9 Saß AC, Lange C, Finger JD, Allen J, Born S, Hoebel J, et al. „Gesundheit in Deutschland aktuell“ – Neue Daten für Deutschland und Europa. Hintergrund und Studienmethodik von GEDA2014/2015-EHIS. *J Health Monit.* 2017;2(1):83-90. doi: 10.17886/RKI-GBE-2017-012.
- 10 Lange C, Finger JD, Allen J, Born S, Hoebel J, Kuhnert R, et al. Implementation of the European health interview survey (EHIS) into the German health update (GEDA). *Arch Public Health* 2017;75:40. doi: 10.1186/s13690-017-0208-6.
- 11 Allen J, Born S, Damerow S, Kuhnert R, Lemcke J, Müller A, et al. Gesundheit in Deutschland aktuell (GEDA 2019/2020-EHIS) – Hintergrund und Methodik. *J Health Monit.* 2021;6(3):72-87. doi: 10.25646/8558.
- 12 Starker A, Schienkiewitz A, Damerow S, Kuhnert R. Verbreitung von Adipositas und Rauchen bei Erwachsenen in Deutschland – Entwicklung von 2003 bis 2023. *J Health Monit.* 2025;10(1):e12990. doi: 10.25646/12990.
- 13 Robert Koch-Institut. Studie „Gesundheit in Deutschland“. 2025 [cited 10.11.2025]. Available from: <https://www.rki.de/DE/Themen/Forschung-und-Forschungsdaten/Sentinel-Surveillance-Panel/Panel/panel-gesundheit-in-deutschland-node.html>.
- 14 Lemcke J, Loss J, Allen J, Öztürk I, Hintze M, Damerow S, et al. Gesundheit in Deutschland: Etablierung eines bevölkerungsbezogenen Gesundheitspanels. *J Health Monit.* 2024;9(S2): 2–22. doi: 10.25646/11913.2.
- 15 Lemcke J, Öztürk I, Damerow S, Heller T, Born S, Wetzstein M, et al. Recruitment of a probability-based general population health panel for public health research in Germany: The panel ‘Health in Germany’. *BMC Med Res Methodol.* Under review.
- 16 Damerow S, Kuhnert R, Schaffrath Rosario A, Lemcke J. Weighting Strategy and Selection Analysis in the Panel ‘Health in Germany’: Methods and Results for the 2024 Annual Survey. *BMC Med Res Methodol.* Forthcoming.
- 17 Lemcke J, Damerow S, Kuttig T, Öztürk I, Albrecht S, Heller T, et al. Panel „Gesundheit in Deutschland“ des Robert Koch-Instituts – Hintergrund und Methode der Jahreserhebung 2024. *J Health Monit.* 2025;10(4):e13546. doi: 10.25646/13546.
- 18 Schroedter JH, Lechert Y, Lüttinger P. Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976-2004 (Version 1). Mannheim: ZUMA-Methodenberichte 2006/08; 2006 [cited 10.11.2025]. Available from: <https://nbn-resolving.org/urn:nbn:de:0168-ss0ar-262397>.
- 19 UNESCO Institute for Statistics. International Standard Classification of Education: ISCED 2011. Montreal, Canada: UIS; 2012.
- 20 Brauns H, Scherer S, Steinmann S. The CASMIN educational classification in international comparative research. In: Hoffmeyer-Zlotnik JHP, Wolf C, editors. *Advances in cross-national comparison*. Boston, MA, USA: Springer; 2003. p. 221-244. doi: 10.1007/978-1-4419-9186-7\_11.
- 21 Hoffmeyer-Zlotnik JHP, Beckmann K, Czajka S, Glemser A, Heckel C, Klumpe B, et al. Demographische Standards: Ausgabe 2024 – Eine gemeinsame Empfehlung des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e. V. (ASI) und des Statistischen Bundesamtes. Köln: GESIS – Leibniz-Institut für Sozialwissenschaften; 2024 [cited 10.11.2025]. doi: 10.21241/ss0ar.94099.
- 22 Behrens K, Böltken F, Dittmar H, Götsche F, Gutfleisch R, Habla H, et al. Regionale Standards: Ausgabe 2019. (3., überarb. u. erw. Aufl.) (GESIS-Schriftenreihe, 23). Köln: GESIS – Leibniz-Institut für Sozialwissenschaften; Arbeitsgruppe Regionale Standards; 2019. doi: 10.21241/ss0ar.6234.
- 23 AAPOR. The American Association for Public Opinion Research. Final Dispositions of Case Codes and Outcome Rates for Surveys. 10th edition. 2023 Standard Definitions [cited 07.11.2025]. Available from: <https://aapor.org/wp-content/uploads/2024/03/Standards-Definitions-10th-edition.pdf>.

**Annex Tabelle 2:** Erwartete Richtung der Methodeneffekte (Unterschiede der Prävalenzschätzer) in Abhängigkeit vom Studiendesign

Indikator	Soziale Erwünschtheit	Faktenfrage (F), Einschätzung (E)	Likert-Skala	Anzahl Items	Erwartete Tendenz	Erläuterung der erwarteten Richtung der Prävalenzunterschiede abhängig vom Studiendesign: RKI-Panel vs. GEDA-Telefon survey
Gute/sehr gute selbst- eingeschätzte allgemeine Gesundheit	gering	E	ja, 5-stufig	1	↘	Der Anteil der Personen mit guter/sehr guter Gesundheit ist im RKI-Panel 2024 etwas niedriger als in GEDA aufgrund des Likert-Skalen-Antwortformats; ggf. ausgeprägter in der Altersgruppe $\geq 80^*$ .
Chronisches Kranksein	gering	F	nein, binär (ja/nein)	1	⇒ ↗	Der Anteil chronisch erkrankter Personen ist im RKI-Panel 2024 höher als in GEDA; ggf. ausgeprägter in der Altersgruppe $\geq 80^*$ .
Einschränkungen bei Alltagsaktivitäten	gering	E + F	ja, 3-stufig	2	⇒ ↗	Der Anteil der Personen, die seit mindestens 6 Monaten mäßig oder stark eingeschränkt sind, unterscheidet sich nicht; ggf. höhere Prävalenz in der Altersgruppe $\geq 80^*$ .
Körperliche Einschränkungen	gering	E	ja, 4-stufig	2	⇒ ↗	Der Anteil der Personen mit großen Schwierigkeiten bzw. Unfähigkeit beim Laufen oder Treppensteigen unterscheidet sich nicht; ggf. höhere Prävalenz in der Altersgruppe $\geq 80^*$ .
Diabetes mellitus	gering	F	nein, binär (ja/nein)	1	⇒ ↗	Der Anteil der Personen mit ärztlich diagnostiziertem Diabetes mellitus unterscheidet sich nicht; ggf. höhere Prävalenz in der Altersgruppe $\geq 80^*$ .
Adipositas	hoch	F	nein, binär (ja/nein)	2	↗	Der Anteil der Personen mit Adipositas (selbstberichteter BMI $\geq 30 \text{ kg/m}^2$ ) ist im RKI-Panel 2024 höher, da der Effekt sozialer Erwünschtheit bei schriftlichen Befragungen geringer ist. Zudem gibt es im Panel die Möglichkeit, Größe und Gewicht nachzumessen und so tendenziell genauere Angaben zu machen.
Rauchen	mittel	F	nein, kategorial	1	↘	Der Anteil der täglich oder gelegentlich rauchenden Personen ist im RKI-Panel niedriger, da sich möglicherweise mehr gesundheitsbewusste Personen an einem Panel zu Gesundheitsthemen beteiligen.
Ausgezeichnete/sehr gute selbsteingeschätzte psychische Gesundheit	mittel	E	ja, 5-stufig	1	↘	Der Anteil der Personen mit ausgezeichneter/sehr guter psychischer Gesundheit ist im RKI-Panel 2024 etwas niedriger als in GEDA aufgrund des Likert-Skalen-Antwortformats oder durch Selektionseffekte**.
Depressive Symptomatik	hoch	E	ja, 4-stufig	8	↗	Der Anteil der Personen mit depressiver Symptomatik (PHQ-8 $\geq 10$ ) ist im RKI-Panel 2024 höher als in GEDA v. a. aufgrund des mehrfachen Likert-Skalen-Antwortformats und da der Effekt sozialer Erwünschtheit bei schriftlichen Befragungen geringer ist, ggf. zusätzlich Selektionseffekte**.
Angstsymptome	hoch	E	ja, 4-stufig	2	↗	Der Anteil der Personen mit Angstsymptomen (GAD-2 $\geq 3$ ) ist im RKI-Panel 2024 höher als in GEDA v. a. aufgrund des mehrfachen Likert-Skalen-Antwortformats und da der Effekt sozialer Erwünschtheit bei schriftlichen Befragungen geringer ist, ggf. zusätzlich Selektionseffekte**.
Soziale Unterstützung	mittel	E	ja, 5-stufig	3	↗	Der Anteil der Personen mit starker sozialer Unterstützung (OSSS-3 $\geq 11$ ) ist im RKI-Panel 2024 niedriger als in GEDA v. a. aufgrund mehrfachen Likert-Skalen-Antwortformat und da der Effekt sozialer Erwünschtheit bei schriftlichen Befragungen geringer ist.

Anmerkungen: GEDA=Gesundheit in Deutschland aktuell, BMI=Body Mass Index, PHQ-8=Patient Health Questionnaire-8, GAD-2=Generalized Anxiety Disorder-2, OSSS-3=Oslo-3-Items-Social-Support-Scale

\*Stichprobeneffekte: bei den älteren Personen mehr Personen mit schlechterer Gesundheit im RKI-Panel

\*\*Stichprobeneffekte: Möglicherweise beteiligen sich mehr Personen mit psychischer Belastung an einem Panel zu Gesundheitsthemen, speziell in einem hauptsächlich auf psychische Gesundheit ausgerichteten Fragebogen.