



GESUNDHEITSBERICHTERSTATTUNG DES BUNDES
GEMEINSAM GETRAGEN VON RKI UND DESTATIS

**Methodische Studie zur Durchführung von
Mixed-Mode-Befragungen zur Gesundheit
von Kindern und Jugendlichen
(Pilotstudie KiGGS Welle 2)**

GESUNDHEITSBERICHTERSTATTUNG DES BUNDES
GEMEINSAM GETRAGEN VON RKI UND DESTATIS

**Methodische Studie zur Durchführung von
Mixed-Mode-Befragungen zur Gesundheit
von Kindern und Jugendlichen
(Pilotstudie KiGGS Welle 2)**

Autorinnen und Autoren

Elvira Mauz
Robert Hoffmann
Robin Houben
Laura Krause
Antje Gößwald
Panagiotis Kamtsiuris

Mitwirkende (in alphabetischer Reihenfolge)

Jennifer Allen
Katharina Behrendt
Hans Butschalowsky
Hans Peter Dircks
Andrea Franke
Christine Gärtner
Robert Hoffmann
Heike Hölling
Robin Houben
Franziska Jentsch
Laura Krause
Michael Lange
Elvira Mauz
Max Prohl
Patrick Schmich
Annika Schmidt
Martin Schwark
Victoria Stier
Marika Varga
Matthias Wetzstein
Yvonne Wüst

Projektleitung

Panagiotis Kamtsiuris

Redaktion

Stefanie Seeling
Thomas Ziese

Inhaltsverzeichnis

	Kernaussagen	6
	Zielstellungen und Studiendesign	6
	Studienergebnisse	6
	Schlussfolgerungen	7
1	Einleitung	9
2	Hintergrund	10
2.1	Die KiGGS-Studie als Bestandteil des Gesundheitsmonitorings am Robert Koch-Institut	10
2.2	Die bisherigen KiGGS-Wellen.	11
2.3	KiGGS Welle 2 im Überblick	12
2.4	Herausforderungen in KiGGS Welle 2	12
3	SAQ-Papier und SAQ-Web im Mixed-Mode-Design: Stand der Forschung .	13
3.1	Kann mit einem Mixed-Mode-Design mit SAQ-Papier und SAQ-Web die Responsequote erhöht werden?	14
3.2	Kann mit einem Mixed-Mode-Design mit SAQ-Papier und SAQ-Web die Repräsentativität erhöht werden?	16
3.3	Treten in einem Mixed-Mode-Design mit SAQ-Papier und SAQ-Web Mode-Effekte auf?	17
3.4	Kann mit einem Mixed-Mode-Design SAQ-Papier und SAQ-Web die Datenqualität und Surveyökonomie erhöht werden?	19
3.5	Zusammenfassung und Fazit.	21
4	Die methodische Pilotstudie: Einbettung, Design und Inhalte	21
4.1	Einbindung der Pilotstudie in den Pretest für KiGGS Welle 2.	21
4.2	Zielstellungen der methodischen Pilotstudie	21
4.2.1	Vergleich unterschiedlicher Erhebungsdesigns	21
4.2.2	Vergleich unterschiedlicher Erhebungsmodi.	22
4.3	Studiendesign der methodischen Pilotstudie	22
4.3.1	Stichprobe.	22
4.3.2	Kontrollgruppendesign	22
4.4	Studieninhalte.	24
4.4.1	Allgemeiner Gesundheitszustand von Kindern und Jugendlichen	24
4.4.2	Wichtige Erkrankungen bei Kindern und Jugendlichen	24
4.4.3	Psychische Auffälligkeiten und Beeinträchtigungen bei Kindern und Jugendlichen	25
4.4.4	Gesundheitsbezogene Lebensqualität von Kindern und Jugendlichen	26
4.4.5	Adipositas bei Kindern, Jugendlichen und deren Eltern	26
4.4.6	Tabakkonsum von Kindern, Jugendlichen und deren Eltern.	27
4.4.7	Alkoholkonsum von Kindern und Jugendlichen	27
4.4.8	Körperliche Aktivität und Sport bei Kindern, Jugendlichen und deren Eltern	28

4.4.9	Mediennutzung von Kindern und Jugendlichen	28
4.4.10	Inanspruchnahme ärztlicher Leistungen von Kindern und Jugendlichen	29
4.5	Methodenübergreifende Operationalisierung der Befragungsunterlagen	29
4.6	Erfassung, Übermittlung und Sicherung der Daten	30
5	Ergebnisse zu Response und Repräsentativität	31
5.1	Vorgehensweise der methodischen Auswertungen	31
5.1.1	Festlegung der Falldefinition	31
5.1.2	Fallzahlen	32
5.2	Teilnahme (Response)	32
5.2.1	Teilnahme nach Erhebungsdesign	32
5.2.2	Teilnahme nach Erhebungsmodus innerhalb der Erhebungsdesigns	34
5.2.3	Nichtteilnahme (Nonresponse)	36
5.3	Stichprobenzusammensetzung	36
5.3.1	Stichprobenzusammensetzung der Erhebungsdesigns	36
5.3.2	Stichprobenzusammensetzung der Modegruppen	38
6	Ergebnisse zu Modeunterschieden zwischen SAQ-Papier und SAQ-Web.	41
6.1	Item-Nonresponse und Datenqualität	41
6.1.1	Methodik	41
6.1.2	Ergebnisse.	41
6.2	Prävalenz- und Mittelwertvergleich von Gesundheitsindikatoren.	44
6.2.1	Methodik	44
6.2.2	Ergebnisse.	45
6.2.3	Fazit zu Mode-Effekten zwischen SAQ-Papier und SAQ-Web.	52
7	Bewertung der Surveyökonomie	54
8	Fazit	56
	Literatur	58

Tabellenverzeichnis

Tabelle 1:	Eingesetzte Fragebögen in der methodischen Pilotstudie	30
Tabelle 2:	Falldefinition für die Studienteilnahme	31
Tabelle 3:	Vollständige und unvollständige Fälle nach Altersgruppen	32
Tabelle 4:	Responseberechnung für die methodische Pilotstudie	33
Tabelle 5:	Responsequoten nach Altersgruppen und Erhebungsdesigns	34
Tabelle 6:	Stichprobenzusammensetzung der realisierten Stichproben und der Bruttostichproben nach Erhebungsdesign	37
Tabelle 7:	Soziodemografische Zusammensetzung der realisierten Stichproben nach Erhebungsmodus (getrennt nach Kontrollgruppe und Mixed-Mode-Gruppen)	39
Tabelle 8:	Fehlende Werte für ausgewählte Gesundheitsindikatoren nach Erhebungsmodus.	42

Abbildungsverzeichnis

Abbildung 1:	Studiendesign der KiGGS-Studie.	10
Abbildung 2:	Bedeutung zusätzlicher Teilnehmerinnen und Teilnehmer per alternative Erhebungsmodus für die Stichprobenzusammensetzung	16
Abbildung 3:	Kontrollgruppendesign der methodischen Pilotstudie	23
Abbildung 4:	Studieninhalte der methodischen Pilotstudie	25
Abbildung 5:	Anteil der Erhebungsmodi bei Kindern und Jugendlichen sowie bei den Eltern nach Erhebungsdesign	35
Abbildung 6:	Selbsteingeschätzter allgemeiner Gesundheitszustand bei Kindern und Jugendlichen nach Erhebungsmodus (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	45
Abbildung 7:	Ausgewählte Erkrankungen bei Kindern und Jugendlichen nach Erhebungsmodus (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	46
Abbildung 8:	Psychische Auffälligkeiten und Beeinträchtigungen bei Kindern und Jugendlichen nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	47
Abbildung 9:	Gesundheitsbezogene Lebensqualität bei Kindern und Jugendlichen nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	48
Abbildung 10:	Adipositas bei Kindern, Jugendlichen und deren Eltern nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	50
Abbildung 11:	Tabakkonsum von Kindern, Jugendlichen und deren Eltern nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	51
Abbildung 12:	Alkoholkonsum von Kindern und Jugendlichen nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	51
Abbildung 13:	Körperliche Aktivität und Sport bei Kindern, Jugendlichen und Eltern nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	52
Abbildung 14:	Mediennutzung von Kindern und Jugendlichen nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	53
Abbildung 15:	Inanspruchnahme ärztlicher Leistungen bei Kindern und Jugendlichen nach Erhebungsmodi (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)	53
Abbildung 16:	Überblick über die zugrundeliegenden Datensätze	55

Kernaussagen

Zielstellungen und Studiendesign

- ▶ Mit der methodischen Pilotstudie KiGGS Welle 2 wurde geprüft, ob durch den Einsatz eines Mixed-Mode-Designs im Vergleich zu einem Single-Mode-Design die Teilnahmebereitschaft steigt, die Verzerrung der Stichprobe verringert wird und eine Vergleichbarkeit der Befragungsergebnisse (Datenqualität) gegeben ist. Eine Rolle spielten auch die Arbeitsabläufe und der zu betreibende Aufwand im Vergleich der verschiedenen Designs.
- ▶ Im Single-Mode-Design wurden schriftliche Fragebögen zum Selbstaussfüllen eingesetzt, während im Mixed-Mode-Design zusätzlich Online-Fragebögen zum Selbstaussfüllen angeboten wurden. Die Einladung zur Studienteilnahme erfolgte stets postalisch.
- ▶ Es wurden verschiedene Mixed-Mode-Erhebungsdesigns angewendet und verglichen: (1) ein sequenzielles Mixed-Mode-Design (Zusenden des Zugangscodes für den Online-Fragebogen mit der Einladung und des schriftlichen Fragebogens erst mit dem Erinnerungsschreiben), (2) ein simultanes Mixed-Mode-Design (gleichzeitiges Zusenden des Online-Zugangscodes und des schriftlichen Fragebogens mit dem Einladungsschreiben) und (3) ein Preselect Mixed-Mode-Design (Anbieten beider Zugangsmöglichkeiten mit dem Einladungsschreiben und eines Antwortformulars zur Anforderung der gewünschten Befragungsunterlagen).

Studienergebnisse

- ▶ In keinem der getesteten Mixed-Mode-Designs konnte eine Erhöhung der Responsequote gegenüber dem Single-Mode-Design mit ausschließlich schriftlicher Befragung und einer Responsequote von 42,0% erreicht werden. Im simultanen Mixed-Mode-Design wird eine Responsequote von 41,4%, im sequenziellen Verfahren von 38,5% und im Preselect-Verfahren von 30,7% erzielt.
- ▶ Aufgrund des Anteils an online ausgefüllten Gesundheitsfragebögen der Eltern von 42,9% im sequenziellen Mixed-Mode-Design, 11,5% im simultanen Mixed-Mode-Design und 33,4%

im Preselect-Design kann die Web-Befragungsform zwar als akzeptiert angesehen werden, jedoch wird die schriftliche Befragungsform bevorzugt. Bei der Befragung der Kinder und Jugendlichen, die ab dem elften Lebensjahr erfolgte, zeigt sich ein ähnliches Bild.

- ▶ Hinsichtlich soziodemografischer Eigenschaften der Teilnehmerinnen und Teilnehmer (Altersgruppen, Gemeindegröße, Region, erste Staatsangehörigkeit, Migrationsstatus, Bildungsstatus) zeigen sich keine Unterschiede zwischen dem Single-Mode-Design und den verschiedenen Mixed-Mode-Designs.
- ▶ Betrachtet man die Teilnehmenden der jeweiligen Befragungsformen SAQ-Papier und SAQ-Web insgesamt, unterscheiden sich die soziodemografischen Eigenschaften sowohl bei den Eltern als auch bei den Kindern und Jugendlichen. Online teilnehmende Eltern haben eine höhere Bildung und ein höheres Einkommen als Eltern, die die Papierform nutzen. Außerdem haben online teilnehmende Eltern jüngere Kinder, sind häufiger verheiratet und leben häufiger in den alten als in den neuen Bundesländern. Kinder und Jugendliche, welche die Web-Befragungsform nutzen, sind eher männlich, streben einen höheren Schulabschluss an und leben in einem Haushalt mit höherem Einkommen und Bildungsstatus.
- ▶ Insgesamt ist nur eine geringe Item-Nonresponse für die beiden angebotenen Erhebungsmodi erkennbar, wobei der schriftliche Fragebogen eine höhere Item-Nonresponse als der Online-Fragebogen aufweist.
- ▶ Die Auswertungen der inhaltlichen Ergebnisse zeigen, dass das Risiko für Mode-Effekte zwischen dem schriftlichen Fragebogen und dem Online-Fragebogen gering ist.
- ▶ Durch Angebot und Nutzung zweier Befragungsformen wurden alle Verwaltungsprozesse in zweifacher Systematik benötigt und alle Fragebögen in zwei Versionen implementiert. Nach der Erhebung lagen zu allen Fragebögen zwei Datensätze vor. Ein solcher Synchronisierungs- bzw. Homogenisierungsaufwand stellt sich in einer Single-Mode-Erhebung nicht. Für sich betrachtet erfordert die webbasierte Erhebungsform eine reduzierte Dateneingabe und -aufbereitung im Vergleich zur schriftlichen Befragung.

Schlussfolgerungen

- ▶ Auf Basis der eingeführten Bewertungskriterien Response, Zusammensetzung der Stichprobe und Datenqualität sind die getesteten Erhebungsdesigns (außer Mixed-Mode Preselect) als gleichwertig anzusehen.
- ▶ Ein höherer Anteil Onlineteilnehmender kann erreicht werden, wenn zuerst nur der Fragebogen in einem sequenziellen Mixed-Mode-Design angeboten wird anstatt beide Modi im simultanen Mixed-Mode-Design gleichzeitig anzubieten.
- ▶ Die Kombination von Papier- und Online-Fragebögen in einem Mixed-Mode-Design ist mit Blick auf Modeffekte vertretbar.
- ▶ Surveyökonomische Aspekte – wie finanzieller, personeller und zeitlicher Aufwand für die Entwicklung einer IT-Infrastruktur, Programmierung, Verwaltungsprozesse, Dateneingabe und Datenverarbeitung – müssen sorgfältig geprüft werden, um für eine konkret geplante Studie das passende Design auszuwählen. Angesichts der hohen Komplexität des Studiendesigns in der Hauptphase von KiGGS Welle 2 – mit insgesamt 13 zu befragenden Untergruppen in Querschnitt und Kohorte, in denen 18 verschiedene Fragebögen eingesetzt werden – wurde der Aufwand für eine Mixed-Mode-Erhebung als unverhältnismäßig hoch bewertet, sodass für KiGGS Welle 2 die Entscheidung für eine ausschließlich schriftliche Befragung getroffen wurde.

Methodische Studie zur Durchführung von Mixed-Mode-Befragungen zur Gesundheit von Kindern und Jugendlichen (Pilotstudie KiGGS Welle 2)

1 Einleitung

Die „Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland“ (KiGGS) ist Bestandteil des Gesundheitsmonitorings für Deutschland, welches das Robert Koch-Institut im Auftrag des Bundesministeriums für Gesundheit durchführt [1, 2]. Im Rahmen des Gesundheitsmonitorings werden zuverlässige Informationen über den Gesundheitszustand, das Gesundheitsverhalten und die gesundheitliche Versorgung der Bevölkerung bereitgestellt. Die KiGGS-Studie hat das Ziel, Daten sowohl für repräsentative Querschnitt- und Trendanalysen für die Altersgruppe der 0- bis 17-jährigen in Deutschland gemeldeten Kinder und Jugendlichen als auch für Längsschnittdaten zu liefern [3]. Die Daten werden durch Befragungen sowie – in bestimmten Zeitabständen – durch körperliche Untersuchungen, Tests und Laboranalysen von Blut- und Urinproben gewonnen [4].

Die KiGGS-Basiserhebung wurde in den Jahren zwischen 2003 und 2006 als kombinierter Untersuchungs- und Befragungssurvey durchgeführt [5–7]. Die Befragung der Eltern aller Teilnehmerinnen und Teilnehmer und der 11- bis 17-jährigen Kinder und Jugendlichen selbst erfolgte mittels schriftlicher Selbstausfüll-Fragebögen (Self-Administered Questionnaires in Papierform, im Folgenden bezeichnet als „SAQ-Papier“). Die erste Folgebefragung KiGGS Welle 1 fand telefonisch im Zeitraum von 2009 bis 2012 statt. Die zweite Folgeerhebung KiGGS Welle 2, deren Feldphase im September 2014 begonnen hat, wird analog zur Basiserhebung sowohl als Untersuchungs- als auch als Befragungssurvey realisiert. Zusätzlich zu den Teilnehmerinnen und Teilnehmern der Basiserhebung wird eine neue Stichprobe über das gesamte Altersspektrum von 0 bis 17 Jahren zur Teilnahme eingeladen. Damit befinden sich in KiGGS Welle 2 zwei voneinander unabhängige Stichproben unter dem Dach der KiGGS-Studie: die neue KiGGS-Querschnitt-

stichprobe sowie die aus der Basiserhebung weitergeführte KiGGS-Kohorte.

Um der zunehmenden Komplexität der Studie Rechnung zu tragen, wurde im Vorfeld von KiGGS Welle 2 die im Folgenden dargestellte methodische Pilotstudie durchgeführt. Ziel dieser Pilotstudie war es zu prüfen, ob die Implementierung von Online-Fragebögen zusätzlich zu schriftlichen Fragebögen sinnvoll ist. Konkret wurde untersucht, welche Auswirkungen auf die Teilnehmerzahl, Stichprobenverzerrung und die Befragungsergebnisse bestehen und wie diese Erkenntnisse unter Effizienzaspekten zu bewerten sind. Anders als in einem Single-Mode-Design, bei dem lediglich ein Erhebungsmedium zum Einsatz kommt, wurden den eingeladenen Personen in verschiedenen Mixed-Mode-Designs neben den schriftlichen Fragebögen auch Online-Fragebögen angeboten. Während die Einladung zur Studienteilnahme in allen Erhebungsdesigns postalisch erfolgte, unterscheiden sich die Mixed-Mode-Designs hinsichtlich der Angebotsabfolgen der beiden Befragungssurveyformen. In einem sequenziellen Mixed-Mode-Design wurde zunächst nur der Zugang zu den Online-Fragebögen und erst zu einem späteren Zeitpunkt eine schriftliche Teilnahmemöglichkeit angeboten. In einem simultanen Mixed-Mode-Design wurden Online-Zugangscodes und schriftlicher Fragebogen mit dem Einladungsschreiben zugesandt. Im sogenannten Preselect Mixed-Mode-Design wurden zwar beide Zugangsmöglichkeiten mit dem Einladungsschreiben ebenfalls parallel angeboten, die eingeladenen Personen hatten die entsprechenden Befragungsunterlagen aber noch anzufordern. Auf Grundlage der Ergebnisse sollte letztlich eine empirisch fundierte Entscheidung für das Erhebungsdesign der Hauptphase von KiGGS Welle 2 getroffen werden.

2 Hintergrund

2.1 Die KiGGS-Studie als Bestandteil des Gesundheitsmonitorings am Robert Koch-Institut

Am Robert Koch-Institut (RKI) werden im Rahmen des Gesundheitsmonitorings in regelmäßigen Abständen bundesweit repräsentative Studien zur gesundheitlichen Lage der Bevölkerung in Deutschland durchgeführt [8]. In den Studien des Gesundheitsmonitorings werden sowohl Kinder und Jugendliche als auch Erwachsene in verschiedenen Studiendesigns untersucht. Eines der Ziele des Gesundheitsmonitorings ist es, die gesundheitliche Lage der in Deutschland lebenden Bevölkerung über alle Lebensphasen darstellen zu können. Die verschiedenen Studien am RKI sind aufeinander abgestimmt und ergänzen sich gegenseitig.

Die KiGGS-Studie untersucht die Gesundheit von Kindern und Jugendlichen und verfolgt

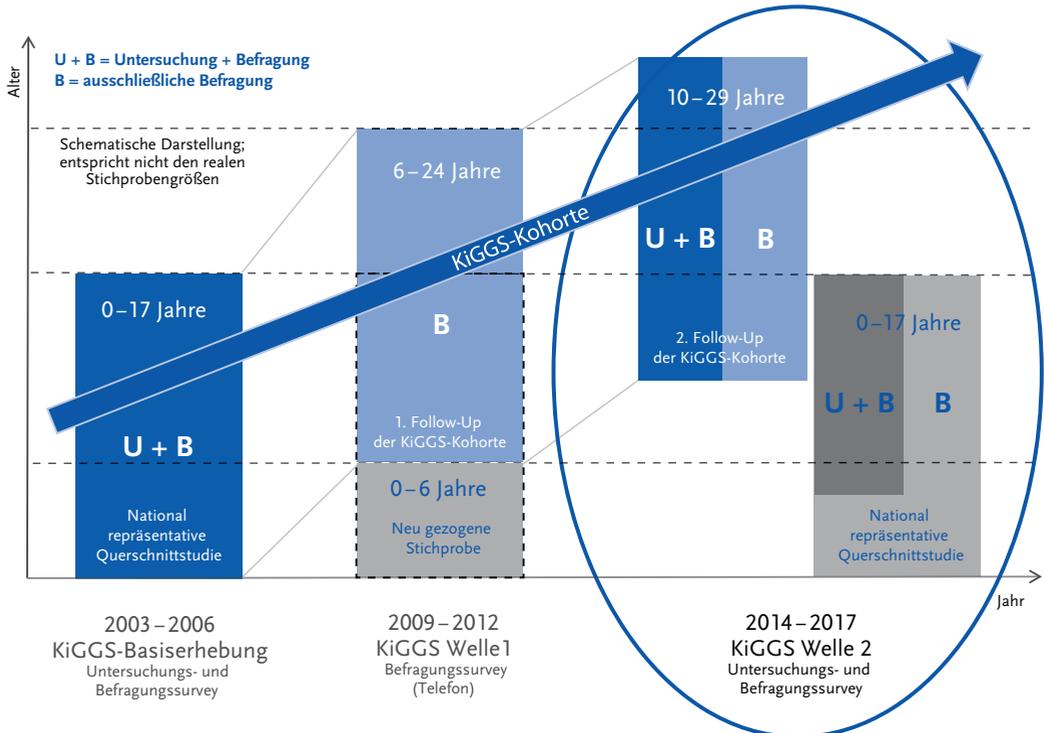
sowohl einen querschnittlichen als auch längsschnittlichen Ansatz (Abbildung 1).

Mit der KiGGS-Studie werden kontinuierliche Datenerhebungen zur gesundheitlichen Lage der Kinder und Jugendlichen in Deutschland vorgenommen [3, 9]. Die Basiserhebung fand zwischen 2003 und 2006 statt, seit 2009 wird KiGGS als Langzeitstudie fortgeführt (Abbildung 1).

Der besondere Wert der KiGGS-Daten liegt in drei zentralen Merkmalen, die in dieser Kombination in Deutschland einzigartig sind:

- (1) KiGGS stellt wiederholt bundesweit repräsentative Daten zur Verfügung, mit denen die aktuelle gesundheitliche Lage der Kinder und Jugendlichen im Alter von 0 bis 17 Jahren sowie zeitliche Entwicklungstrends beschrieben werden können. Bereits 2014 wurden mit Daten der KiGGS Welle 1 für eine Vielzahl von Gesundheitsindikatoren Analysen zur zeitlichen Entwicklung im Vergleich zur Basiserhebung veröffentlicht.

Abbildung 1
Studiendesign der KiGGS-Studie



- (2) KiGGS beinhaltet eine Kohorte, d. h. Mädchen und Jungen der Basiserhebung werden bis in das Erwachsenenalter wiederholt eingeladen. Dadurch können Ursachen und Einflussfaktoren gesundheitlicher Entwicklungen im Lebensverlauf analysiert werden. In KiGGS Welle 1 waren die Teilnehmerinnen und Teilnehmer dabei 6 bis 24 Jahre alt. Mit Abschluss von KiGGS Welle 2 werden Analysen über zwei bis drei Messzeitpunkte möglich und die Teilnehmerinnen und Teilnehmer 10 bis 29 Jahre alt sein.
- (3) In KiGGS werden die Daten mit Hilfe von Befragungen sowie medizinischen Untersuchungen, Tests und Laboranalysen erhoben. Dies erlaubt gegenüber reinen Befragungssurveys validere Aussagen auf der Basis messbarer Parameter wie z. B. bei Körpergröße und -gewicht und verbesserte Häufigkeitsschätzungen für Krankheiten und Beschwerden.

Die Daten der KiGGS-Studie liefern gemeinsam mit anderen Datenquellen eine umfassende Daten- und Informationsgrundlage für die Akteure in der Public-Health-Landschaft und die Gesundheitspolitik in Deutschland. Durch das Identifizieren von Gesundheitsrisiken im Kindes- und Jugendalter leisten die Daten einen Beitrag zur Entwicklung geeigneter Präventionskonzepte und -maßnahmen. Die Ergebnisse der KiGGS-Studie fließen in die Publikationen der Gesundheitsberichterstattung (GBE) des Bundes am RKI sowie in zahlreiche Forschungsarbeiten ein [5, 9–11].

2.2 Die bisherigen KiGGS-Wellen

Die Basiserhebung wurde zwischen 2003 und 2006 durchgeführt. In zufällig ausgewählten 167 Orten (Sample Points) haben insgesamt 17.641 Kinder und Jugendliche im Alter von 0 bis 17 Jahren (8.656 Mädchen und 8.985 Jungen) sowie deren Eltern an der Studie teilgenommen (Responsequote von 66,6 %) [6]. Das Studienprogramm umfasste verschiedene körperliche Untersuchungen (einschließlich Laboranalysen von Blut- und Urinproben) und Tests für die Kinder und Jugendlichen sowie das Ausfüllen schriftlicher Fragebögen durch die Eltern und Kinder und Jugendlichen selbst ab dem elften Lebensjahr. Inhaltliche

Schwerpunkte der Basiserhebung zielten auf den Gesundheitszustand, das Gesundheitsverhalten, die Lebensbedingungen, Schutz- und Risikofaktoren und die Inanspruchnahme von Leistungen des Gesundheitssystems durch Kinder und Jugendliche in Deutschland. Ergänzt wurde die Basiserhebung durch Module, die an Unterstichproben die Beantwortung spezifischer Fragestellungen in Bezug auf die Themen „Psychische Gesundheit“ (BELLA), „Motorik“ (MoMo), „Umwelt“ (Kinder-Umwelt-Survey (KUS)) und „Ernährung“ (EsKiMo) erlaubte [6, 7, 12, 13].

Die im Zeitraum von 2009 bis 2012 durchgeführte KiGGS Welle 1 ist die erste Folgeerhebung der KiGGS-Studie. Anders als in der Basiserhebung erfolgten in KiGGS Welle 1 ausschließlich telefonische Befragungen (Untersuchungen wurden nicht durchgeführt). Die Teilnehmerinnen und Teilnehmer der Basiserhebung, die zu diesem Zeitpunkt zwischen 6 und 24 Jahre alt waren, wurden hierfür erneut eingeladen. Gleichzeitig wurden für den Altersbereich von 0 bis 6 Jahren neue Teilnehmerinnen und Teilnehmer in die Studie einbezogen, die erneut aus den im Rahmen der Basiserhebung gezogenen 167 Gemeinden zufällig ausgewählt wurden. Insgesamt nahmen 12.368 Kinder und Jugendliche (6.093 Mädchen und 6.275 Jungen) in dem für den Querschnitt relevanten Altersbereich von 0 bis 17 Jahren teil; davon 4.455 Ersteingeladene zwischen 0 und 6 Jahren (Responsequote: 38,8 %) und 7.913 Wiedereingeladene zwischen 7 und 17 Jahren (Responsequote: 72,9 %) [4]. Im Altersbereich von 18 bis 24 Jahren haben insgesamt 3.491 junge Erwachsene (1.852 Frauen und 1.639 Männer) erneut an der Studie teilgenommen (Responsequote: 59,9 %) [14]. Mit KiGGS Welle 1 konnten erneut aktuelle, bundesweit repräsentative Daten zur gesundheitlichen Lage der 0- bis 17-jährigen Kinder und Jugendlichen in Deutschland gewonnen werden. Inhaltliche Schwerpunkte von KiGGS Welle 1 umfassten die körperliche und psychische Gesundheit, das Gesundheitsverhalten und die soziale Lage. Als Zusatzmodule wurden die vertiefenden Module der Basiserhebung zu den Themen „Psychische Gesundheit“ (BELLA) und „Motorik“ (MoMo) fortgesetzt. Ferner fanden im Rahmen einer landesweiten Stichprobenaufstockung repräsentative Datenerhebungen für das Land Thüringen (Landesmodul Thüringen) statt [1, 4, 9, 15].

2.3 KiGGS Welle 2 im Überblick

Die KiGGS Welle 2, die zwischen 2014 und 2017 durchgeführt wird, wird analog zur Basiserhebung sowohl als Befragungs- als auch als Untersuchungssurvey realisiert. Wie eingangs erwähnt befinden sich in KiGGS Welle 2 zwei voneinander unabhängige Stichproben mit unterschiedlichen Zielstellungen und Verfahrensweisen unter dem Dach der Studie:

(1) Im Einklang mit den vorherigen KiGGS-Erhebungen soll auch KiGGS Welle 2 aktuelle, bundesweit repräsentative Daten zur gesundheitlichen Situation der 0- bis 17-jährigen Mädchen und Jungen mit Wohnsitz in Deutschland liefern und zugleich Trendanalysen über nunmehr drei Zeitpunkte erlauben. An den 167 Untersuchungsstellen der Basiserhebung wird hierzu eine neue Einwohnermeldeamtsstichprobe gezogen. Angestrebt wird eine realisierte Stichprobe von ungefähr 12.750 Teilnehmerinnen und Teilnehmern. Während ein Teil der Zielpersonen ab 3 Jahren sowohl zur Befragung als auch zur Untersuchung in ein Untersuchungszentrum eingeladen wird, erhalten die übrigen Kinder und Jugendlichen beziehungsweise deren Eltern nur eine Einladung für den Befragungsteil. Die Einteilung in diese Gruppen erfolgt zufällig. Zudem wird KiGGS Welle 2 erneut durch verschiedene Modulstudien ergänzt. Diese zielen thematisch auf die „Psychische Gesundheit“ (BELLA), die „Motorik“ (MoMo), die „Umwelt“ (GerES) und die „Ernährung“ (KiESEL-Studie für Kinder im Alter von 6 Monaten bis 5 Jahren, EsKiMo für Kinder und Jugendliche im Alter von 6 bis 17 Jahren) [3, 16].

(2) In der KiGGS-Kohorte werden von den 17.641 Teilnehmerinnen und Teilnehmern der Basiserhebung, unabhängig von ihrer Teilnahme an KiGGS Welle 1, diejenigen erneut eingeladen, die einer Wiederkontakttierung zugestimmt hatten, nicht ins Ausland verzogen oder verstorben sind.

2.4 Herausforderungen in KiGGS Welle 2

Die doppelte Ausdifferenzierung des Stichprobendesigns im Vergleich zur KiGGS-Basiserhebung in eine quer- und längsschnittliche Komponente sowie die jeweils möglichen unterschiedlichen Erhebungstypen (Befragung und Untersuchung

oder nur Befragung) stellen anspruchsvolle Rahmenbedingungen an die Durchführung von KiGGS Welle 2.

Maßgeblich ist dabei vor allem, dass sowohl für die bevölkerungsbezogenen Querschnittanalysen als auch die Lebensverlaufsanalysen hohe Teilnehmerzahlen angestrebt werden. Mit Blick auf erstere reduzieren diese die Wahrscheinlichkeit einer systematischen Verzerrung der Stichprobe und erhöhen die statistische Power auch bei niedrigprävalenten Erkrankungen. Die Implementierung verschiedener Erhebungstypen kann zur Erreichung dieser Ziele beitragen, bringt jedoch auch neue organisatorische und logistische Anforderungen mit sich.

Zentrales Ziel der Pilotstudie war die Ermittlung von Maßnahmen zur Optimierung der Datenerhebung, um eine hohe Teilnehmerzahl bei bevölkerungsrepräsentativer Zusammensetzung der realisierten Stichprobe zu erreichen. Zudem standen eine hohe Qualität und Vergleichbarkeit der inhaltlichen Daten im Vordergrund. Ein wesentliches Kriterium war zudem, dass die Datenerhebung effizient durchgeführt werden kann und dass die Daten nach Ende der Erhebung schnell für die Auswertung zur Verfügung stehen. Damit sollte der geschilderte hohe planerische Aufwand im Vorfeld der Studie möglicherweise kompensiert werden können.

Als eine vielversprechende Maßnahme wurde die Implementierung eines webbasierten Selbstausfüll-Fragebogens (SAQ-Web – Self-Administered Questionnaire mit Web-Fragebogen) erwogen – als Weiterentwicklung des bisherigen Single-Mode-Designs der KiGGS-Studie mit dem Einsatz ausschließlich schriftlicher Fragebögen (SAQ-Papier) in ein Mixed-Mode-Design. Ein solches wird am RKI bereits für die regelmäßige Gesundheitsbefragung der Erwachsenenbevölkerung (GEDA-Studie) angewendet [17]. Allerdings lassen sich die bereits erzielten Ergebnisse für die Erwachsenenbevölkerung ohne eigene methodische Testung nicht auf die Studienpopulation der KiGGS-Studie übertragen. Die bereits zur Verfügung stehende Overhead-Infrastruktur des RKI für die Durchführung von Web-Befragungen bot jedoch die Möglichkeit, eine solche Testung in der vorliegenden methodischen Pilotstudie ohne diesbezügliche Anschaffungskosten vorzunehmen.

Als mögliche Vorteile der zusätzlich angebotenen Web-Befragungsform im Rahmen eines

Mixed-Mode-Designs wurden die automatisierte Filterführung, die direkte Dateneingabe durch die Teilnehmenden selbst (Wegfall personalintensiver manueller Dateneingaben, Qualitätssicherung während der Dateneingabe) sowie reduzierte Druck- und Portokosten gesehen. Maßgeblich war aber insbesondere die Überlegung, dass durch das Angebot eines zusätzlichen Erhebungsmodus Personen für die Teilnahme gewonnen werden könnten, die die Teilnahme ansonsten abgelehnt hätten. In einer vermutlich internetaffinen Studienpopulation von Jugendlichen, jungen Erwachsenen sowie jüngeren Eltern könnte dies besonders zu erwarten sein. Neben einer höheren Ausschöpfung der Bruttostichprobe könnte sich hieraus auch eine verbesserte Zusammensetzung der realisierten Stichprobe im Vergleich zur bereinigten Bruttostichprobe ergeben, sollten sich die Nutzerinnen und Nutzer der webbasierten und schriftlichen Fragebögen systematisch voneinander unterscheiden.

Diesen Potenzialen einer zusätzlich angebotenen Teilnahmeform steht jedoch das Risiko einer sinkenden Vergleichbarkeit der inhaltlichen Daten gegenüber. Sogenannte Mode-Effekte können eine systematische Fehlerquelle bilden, die zu verzerrten und nicht vergleichbaren Ergebnissen führen kann [18]. Dies ist gerade für Studien von Bedeutung, die Trendaussagen machen und individuelle Entwicklungsverläufe aufzeigen. Zudem sind etwaige Mehraufwände in der Konzeption und Durchführung des Surveys zu erwägen. Zu nennen sind hier insbesondere die Diversifikation der Verwaltungsprozesse, die Bereitstellung der Fragebögen in Papier- und Onlineform sowie die notwendige Homogenisierung der Erhebungsdaten aus diesen Quellen.

Ausgangspunkt für die Planungen von KiGGS Welle 2 sind zusammenfassend die Fragestellungen, ob mit der Aufnahme des webbasierten Erhebungsmodus (SAQ-Web) in ein Mixed-Mode-Design im Vergleich zur Single-Mode-Erhebung (SAQ-Papier), wie sie in der Basiserhebung implementiert war,

- (1) die Response erhöht,
- (2) die Stichprobenzusammensetzung verbessert,
- (3) vergleichbare Ergebnisse erreicht und
- (4) die Datenqualität und Surveyökonomie verbessert werden können.

Im Folgenden wird mit Blick auf diese vier Fragestellungen der Stand der Forschung zum Zeitpunkt der Pilotstudie zur Einbindung der beiden Erhebungsformen in verschiedene Mixed-Mode-Designs beschrieben.

3 SAQ-Papier und SAQ-Web im Mixed-Mode-Design: Stand der Forschung

Bei der Planung einer Studie steht am Anfang idealerweise die Frage, wie mit einem limitierten Budget durch ein optimales Surveydesign eine maximale Datenqualität im Sinne niedriger systematischer Fehler erreicht werden kann. Der zugrundeliegende „Total Survey Error“-Ansatz [19, 20] liefert eine umfassende Konzeptualisierung potenzieller Fehlerquellen im gesamten Studienverlauf unter Einbeziehung einer kontextbezogenen Perspektive. Der Total Survey Error setzt sich aus dem Sampling- und dem Nonsampling-Error zusammen [20]. Um einen geringen Sampling-Error zu erreichen, sollte das Stichprobendesign, die zu erreichende realisierte Stichprobengröße sowie der zu berechnende Schätzer in der Form gewählt werden, dass die Fragestellung mit einer hinreichenden statistischen Power beantwortet werden kann. Der Hauptfokus bei der Konzeption des Studiendesigns liegt auf dem Nonsampling-Error [21], der sich ebenfalls wiederum aus verschiedenen Fehlerquellen zusammensetzt. Dazu gehören (1) die sogenannte Spezifikation, d. h. die Präzision des Surveys, (2) Frame- und Coverage-Error, d. h. Fehler im Zusammenhang mit der Bearbeitung des Stichprobenrahmens, (3) der Nonresponse-Error, der aus Fehlern aus der systematischen Nichtteilnahme verschiedener Zielpopulationsgruppen oder der Item-Nonresponse resultiert, (4) Messfehler, die von den Interviewenden, den Befragten oder der Erhebungsmethode selbst ausgehen können und (5) Fehler im Kontext des Datenmanagements [20–22]. Ziel bei der Auswahl eines Studiendesigns ist, das Zusammenspiel potenzieller Fehlerquellen zu berücksichtigen und damit den Gesamtfehler so niedrig wie möglich zu halten.

Die seit den 1990er-Jahren kontinuierlich rückläufigen Responsequoten in epidemiologischen Studien [21, 23, 24] erhöhen die Gefahren, welche von diesen Fehlerquellen ausgehen. Besonders

problematisch ist diese Entwicklung in schwer zu erreichenden Bevölkerungsgruppen, beispielsweise bei sehr mobilen Gruppen wie jungen Erwachsenen [25], Vollzeitbeschäftigten, Menschen mit Migrationshintergrund oder Bewohnerinnen und Bewohnern von Großstädten [26]. Um dieser Entwicklung entgegenzusteuern und ein hinreichend großes und repräsentatives, das heißt entsprechend den Eigenschaften der Brutstichprobe zusammengesetztes Sample zu erreichen, müssen immer höhere Kosten einkalkuliert werden [27].

Zur Reduzierung einer oder mehrerer im Survey enthaltener systematischer Fehlerquellen unter Berücksichtigung finanzieller und zeitlicher Restriktionen werden Mixed-Mode-Designs zunehmend diskutiert [u. a. 22, 24, 25, 28, 29–31]. In einem auf die Datenerhebung bezogenen Mixed-Mode-Design sollen die Vorteile mehrerer Erhebungsmodi durch ihren gleichzeitigen Einsatz genutzt, gleichzeitig aber deren jeweilige mögliche Defizite ausgeglichen werden [22]. Mit Blick auf die Verringerung des Total Survey Errors werden im Einzelnen vor allem die folgenden Vorteile diskutiert [u. a. 31, 32, 33–36]:

- (1) eine Erhöhung der Response aufgrund der angebotenen Wahlmöglichkeiten,
- (2) eine Reduzierung des Coverage- und Nonresponse-Errors dadurch, dass mit verschiedenen Erhebungsmodi möglicherweise unterschiedliche Teile der Zielpopulation erreicht und für die Teilnahme gewonnen werden können sowie
- (3) eine Reduzierung der allgemeinen Kosten durch den Einsatz günstiger Erhebungsmodi.

Allerdings geht nicht notwendigerweise mit einem Mixed-Mode-Design ein reduzierter Total Survey Error einher, da in einem solchen Design bestimmte systematische Fehler erst ermöglicht werden [22, 33, 35]. Besonders bedeutsam ist die potenzielle Erhöhung des Messfehlers über Mode-Effekte durch den Einsatz verschiedener Erhebungsmethoden [28, 32, 37, 38] oder aufgrund einer höheren Fehleranfälligkeit durch die Notwendigkeit zur Homogenisierung verschiedener Instrumente. Bei der Abwägung von Vor- und Nachteilen ist neben möglichen Fehlerquellen zudem ein erhöhter Aufwand durch Anpassung der Fragebögen und Homogenisie-

rung der Daten aus verschiedenen Quellen [24] zu berücksichtigen.

Neben der grundsätzlichen Entscheidung für oder gegen ein Mixed-Mode-Design muss geklärt werden, in welchem Mixed-Mode-Design die Kombination der verschiedenen Erhebungsmethoden erfolgen soll [24]. In einem sequenziellen Mixed-Mode-Design wird im initialen Kontakt zunächst der eine, meist kostengünstigere Antwortmodus und in den Folgekontakten zusätzlich ein oder mehrere andere Modi angeboten. In einem parallelen oder konkurrenten Mixed-Mode-Design werden zeitgleich alle verfügbaren Erhebungsmodi angeboten [22, 39, 40].

Zentrale Beurteilungskriterien einer solchen Gegenüberstellung von Vor- und Nachteilen sind die Responsequote, eine Verzerrung der Stichprobe zu vermeiden, die Validität und Reliabilität der erhobenen Daten [21] sowie Überlegungen zum erreichten Verhältnis aus möglicherweise gesparten Kosten und erhöhtem Nutzen [24].

3.1 Kann mit einem Mixed-Mode-Design mit SAQ-Papier und SAQ-Web die Responsequote erhöht werden?

Ein zentrales Argument für Mixed-Mode-Designs liegt in der Annahme, eine weitere Erhebungsmethode würde zu zusätzlichen Teilnehmenden führen, die andernfalls nicht teilgenommen hätten [18, 40]. Begründet wird dies über eine Motivationserhöhung durch die Auswahlmöglichkeit im Allgemeinen [35, 40], respektive durch die Möglichkeit, im jeweils präferierten Antwortmodus teilnehmen zu können [40, 41].

Empirische Studien erbringen variierende Ergebnisse in Abhängigkeit von den konkret eingesetzten Methoden und dem Single-Mode-Design, das mit dem Mixed-Mode-Design verglichen wird. Dabei ist vorwegzustellen, dass in Web-Befragungen, die als Single-Mode-Designs konzipiert sind, zumeist deutlich niedrigere Responsequoten als in einem Survey mit ausschließlich schriftlicher Befragung erzielt werden [21, 33, 37, 42–49]. Allgemein erreichen Studien eher selten eine höhere Responsequote in SAQ-Web-Surveys, wofür sie bisher noch keine abschließende Erklärung liefern konnten [33]. Vor allem Studien, die junge und gebildete Erwachsene

betrachten, und Studien mit sensiblen Inhalten berichten widersprüchliche Ergebnisse hinsichtlich der Responsequote [42, 50–53]. Als Schlussfolgerung der höheren Response in schriftlichen Befragungsdesigns wird nahegelegt, in bevölkerungsbezogenen Studien die Befragung mittels SAQ-Web aufgrund der niedrigen erzielten Response ausschließlich in einem Mixed-Mode-Design einzusetzen [43]. Dies ist in der vorliegenden Pilotstudie der Fall.

Mit Blick auf die Kombination der schriftlichen Befragung mit der Online-Erhebung in einem parallelen oder sequenziellen Mixed-Mode-Design berichten diverse Methodenstudien von einer deutlichen Erhöhung der Responsequote im Vergleich zu einem Single-Mode-Design mittels SAQ-Web-Befragung [u. a. 37, 54, 55].

Die empirische Befundlage stimmt allerdings skeptisch hinsichtlich einer Responseerhöhung durch die Methodenkombination, wenn parallele oder sequenzielle Mixed-Mode-Designs mit einem ausschließlich schriftlichen Erhebungsdesign verglichen werden. Zwar finden sich hier ebenfalls höhere Responsequoten in den Mixed-Mode-Designs in verschiedenen Studien, teilweise aber speziellen Studienpopulationen wie bei Collegestudierenden im parallelen Design [56], bei Fachärztinnen und Fachärzten für Kinderheilkunde im sequenziellen Design [57] und bei einer Gruppe sehr gut gebildeter Personen mit hohem technischen Knowhow ebenfalls mit parallelem Mixed-Mode-Design [55]. Aber auch bevölkerungsbezogene Studien berichten von höheren Responsequoten wie der deutsche Epidemiologische Suchtsurvey im sequenziellen Design [58] oder auch der amerikanische Zensus sowohl im parallelen [59] als auch im sequenziellen Design [60]. Einige Studien berichten auch vergleichbare Responsequoten [41, 60–63].

Oftmals werden für die Mixed-Mode-Designs vergleichend zur ausschließlich schriftlichen Befragung jedoch niedrigere Responsequoten erzielt. Medway und Fulton (2012) verglichen in ihrer Meta-Analyse 19 experimentelle Studien, in denen simultane Mixed-Mode-Designs mit SAQ-Papier- und SAQ-Web-Befragung jeweils einer Single-Mode-Erhebung in SAQ-Papier-Form gegenübergestellt wurden. Im Mixed-Mode-Design wird dabei durchschnittlich eine um 3,8 % niedrigere Responsequote festgestellt [40]. In die-

ser Meta-Analyse ist eine mit der Studienpopulation von KiGGS Welle 2 vergleichbare Studie mit Jugendlichen und jungen Erwachsenen im Alter von 16 bis 24 Jahren enthalten, in der im Single-Mode-SAQ-Papier-Design die höchste Responsequote erreicht wird [64]. Zu gleichen Ergebnissen kommt eine Methodenstudie, die 2011 im Rahmen des dänischen Arms der NordChild-Studie durchgeführt wurde. Es fand eine schriftliche Gesundheitsbefragung der Eltern von 2- bis 17-Jährigen in vier unterschiedlichen Erhebungsdesigns statt. In ihrem Vergleich sowohl web- als auch papierbasierten Single-Mode-Designs mit einem Mixed-Mode-Design beider Methoden war die Responsequote in der Single-Mode-Gruppe mit SAQ-Papier-Befragung um 9–11 % höher als in den anderen Gruppen [65]. In einer Meta-Studie zu drei experimentellen Vergleichen mit sequenziellen Mixed-Mode-Designs, bei denen zunächst das Web-Angebot unterbreitet wurde, wird eine um 10–12 % niedrigere Responsequote im Vergleich zum Single-Mode-Design berichtet [66].

Zur Erklärung der Befunde niedrigerer Responsequoten in Mixed-Mode-Designs mit einem zusätzlichen Web-Angebot neben der Möglichkeit, schriftliche Fragebögen zu beantworten, werden verschiedene Ansätze diskutiert:

- (1) Ein zusätzlich angebotener Modus erhöht die Komplexität des Antwortprozesses [40], beispielsweise im Sinne der „Paradox of Choice“-Theorie [64], wonach ein Zuviel an Optionen durch ein Überflutungs- und Überforderungsgefühl einen negativen Effekt erzeugt,
- (2) die Wahl des Web-Angebotes verursacht bei postalischer Einladung einen Medienbruch innerhalb des Antwortprozesses [67],
- (3) der Versuch der Nutzung des Online-Fragebogens unterliegt einem (technischen) Anwendungsfehler und wird dann abgebrochen [40, 49] sowie
- (4) die Datensicherheit bei der Web-Befragung wird als problematisch erachtet [56, 67].

Die Schlussfolgerungen zum Einsatz von Mixed-Mode-Designs mit dem kombinierten Einsatz von SAQ-Web- und SAQ-Papier-Fragebögen sind hinsichtlich der Response dementsprechend nicht einheitlich. Es bedarf so der Betrachtung anderer möglicher Vor- und Nachteile eines solchen Mixed-

Mode-Designs. Ein wichtiges Argument sind die reduzierten Kosten aufgrund der Online-Teilnehmenden, die auch bei geringerer Response ausschlaggebend sein können. Vor diesem Hintergrund betonen Millar und Dillman (2011) die entscheidungsrelevante Bedeutung des erwünschten und erreichten Zuwachses an Web-Teilnehmenden in einem sequenziellen Mixed-Mode-Design, in dem als erstes der Online-Fragebogen angeboten wird. Sie betrachten die Nutzung eines Mixed-Mode-Designs daher lediglich in einem sequenziellen Verfahren als äquivalent hinsichtlich Response und anderen Gesichtspunkten zu einem Single-Mode-Design mit SAQ-Papier-Befragung [41]. Trotz einer um 11% geringeren Responsequote erachten Hohwü et al. (2013) in der Methodenstudie der NordChild-Studie aufgrund deutlich geringerer Kosten die Web-Befragung als geeignet, die Befragung per schriftlichen Fragebogen abzulösen [65]. Im Rahmen einer neuseeländischen Studie an der Allgemeinbevölkerung empfiehlt Brennan (2005) aufgrund einer niedrigeren Responsequote im simultanen Mixed-Mode-Design allerdings die ausschließlich schriftliche Befragung [68]. Zu diesem Schluss kommen auch Turner et al. (2010) in einer Studie mit einer Gruppe von Jugendlichen aufgrund der leicht erhöhten Responsequote und dem gesparten Aufwand für die Programmierung des Web-Fragebogens [64].

Zusammenfassend lässt der Blick auf die Empirie vermuten, dass eine Erhöhung der Responsequote über die Addition eines Online-Modus zur schriftlichen Befragung nicht wahrscheinlich ist. Allerdings handelt es sich bei den Studienpopulationen der KiGGS-Studie, also bei Eltern von 0- bis 17-jährigen Kindern und Jugendlichen und bei den selbst befragten Kindern und Jugendlichen um junge und möglicherweise internetaffine Menschen, die eher Online-Fragebögen wählen könnten.

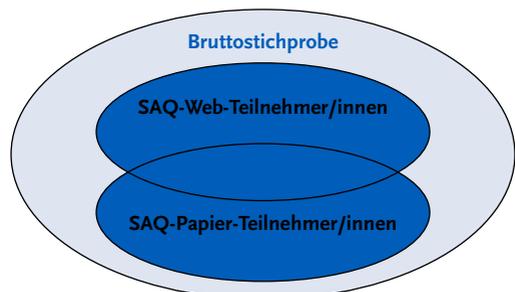
3.2 Kann mit einem Mixed-Mode-Design mit SAQ-Papier und SAQ-Web die Repräsentativität erhöht werden?

Ein weiteres Argument für Mixed-Mode-Designs liegt in der Annahme, dass über das Angebot einer weiteren Erhebungsmethode nicht nur zusätzliche Teilnehmende, sondern auch oder insbesondere Teilnehmerinnen und Teilnehmer mit ande-

ren Eigenschaften gewonnen werden könnten. Dadurch ließe sich die Stichprobenzusammensetzung und damit die Repräsentativität einer Studie gegenüber den jeweiligen einzelnen Erhebungsmodi in einem Single-Mode-Design verbessern [18, 35, 69, 70].

Abbildung 2 zeigt exemplarisch, dass mit der Nutzung schriftlicher Fragebögen ein gewisser Anteil an Eingeladenen aus der Bruttostichprobe für die Teilnahme mit dieser Befragungsform gewonnen werden kann. Werden zudem Web-Fragebögen angeboten, könnten möglicherweise zusätzliche Teilnehmerinnen und Teilnehmer rekrutiert werden, die bei einem ausschließlichen Angebot von schriftlichen Fragebögen nicht zu gewinnen gewesen wären. Zudem gibt es Teilnehmende, die unabhängig vom Erhebungsmodus an der Studie teilgenommen hätten (überlappender Bereich). Ähnlich verhält es sich mit der Stichprobenzusammensetzung. Hier ist die Bruttostichprobe in Abbildung 2 als Ansammlung verschiedener Kombinationen an Eigenschaften zu verstehen. Demnach könnten SAQ-Papier-beziehungsweise SAQ-Web-Teilnehmende eine jeweils für sich „homogene“ Gruppe bilden, die sich aber voneinander unterscheiden. Allerdings ziehen durch eine zusätzliche Befragungsform gewonnene Teilnehmerinnen und Teilnehmer keinesfalls zwingend eine veränderte Komposition der realisierten Stichprobe mit sich, wenn es keine Unterschiede zwischen denjenigen gibt, die entweder mit SAQ-Papier oder mit SAQ-Web geantwortet haben (überlappender Bereich).

Abbildung 2
Bedeutung zusätzlicher Teilnehmerinnen und Teilnehmer per alternativen Erhebungsmodus für die Stichprobenzusammensetzung



SAQ = Self-Administered Questionnaire

Bei einem Angebot von Web-Fragebögen besteht grundsätzlich die Limitierung, dass nur jene Personen teilnehmen können, die Zugang zu einem Internetanschluss haben. Zudem muss eine gewisse Internetaffinität gegeben sein [22, 24, 47, 71]. Dies kann zu einem Coverage-Fehler und/oder Selektionseffekten führen. Beispielsweise zeigt sich in einer Vielzahl von Studien, dass sich die Stichprobenzusammensetzungen im direkten Vergleich der beiden Erhebungsmethoden SAQ-Papier und SAQ-Web systematisch voneinander unterscheiden [18]. Die Präferenz für SAQ-Web ist dabei gemäß eines „Digital Divide Effektes“ [72] über einen höheren sozioökonomischen Status determiniert. Online teilnehmende Personen haben eine höhere Bildung, ein höheres Einkommen sowie eine höhere berufliche Stellung [45, 63, 66, 73–79] und sind jünger [37, 45, 57, 72–74, 77–81]. Der Altersaspekt ist im Hinblick auf die KiGGS-Studie besonders relevant, da unter Jugendlichen die SAQ-Web-Befragung sehr gut als eine Form der Gesundheitsbefragung angenommen wird [82]. Mitunter werden Web-Befragungen daher nur für spezielle Zielpopulationen empfohlen [49, 68] beziehungsweise Mixed-Mode-Designs werden als Möglichkeit gesehen, diese Probleme auszuräumen [22]. Hier gelten die unterschiedlichen Stichprobenzusammensetzungen gleichzeitig als Chance, ein besseres Abbild der Zielpopulation zu erreichen, da beispielsweise die ältere Bevölkerung in erster Linie die schriftliche Befragung nutzt [24].

Die methodische Vorgehensweise von vergleichenden Analysen zur Stichprobenzusammensetzung des kombinierten Einsatzes von SAQ-Papier und SAQ-Web in einem Mixed-Mode-Design mit der ausschließlich schriftlichen Befragung ist der Vergleich der Verteilungen von Stichprobeneigenschaften der jeweiligen realisierten Stichproben untereinander sowie mit den Eigenschaften der bereinigten Bruttostichprobe. Allerdings behindert die Beschränkung auf spezifische Zielpopulationen in empirischen Studien die Übertragbarkeit der Ergebnisse auf bevölkerungsbezogene Studien. Größtenteils wird keine Verbesserung der Stichprobenzusammensetzung festgestellt, beispielsweise auch nicht in einer Befragung von 16- bis 24-Jährigen [64], einer für KiGGS relevanten Zielgruppe, oder in einer Gesundheitsstudie unter Brustkrebspatientinnen [63]. In einer

bevölkerungsbezogenen Studie kommt Brennan (2005) zu dem Schluss, dass zwar ein SAQ-Papier-Angebot die Verzerrung der Stichprobe von Teilnehmerinnen und Teilnehmern an der Web-Befragung kompensieren, aber im Vergleich zu einem Single-Mode-Design mit SAQ-Papier-Befragung nicht verbessern kann [68].

Insgesamt ist die Befundlage zu der hier konkret behandelten Fragestellung quantitativ begrenzt. Die meisten Studien konzentrieren sich auf den Vergleich der Modegruppen, nutzen andere Erhebungsformen, andere experimentelle Erhebungsdesigns oder fokussieren auf die Responsequoten. Die vergleichbar aufgebaute Nord-Child-Studie 2011 in Dänemark berichtet zwar von subgruppenspezifischen Responsequoten, allerdings nicht von Vergleichen der Gruppenzusammensetzungen zwischen den experimentellen Gruppen [65].

Insgesamt kann mit Blick auf die Verbesserungsmöglichkeiten der Stichprobenzusammensetzung durch ein Mixed-Mode-Design vergleichend zur schriftlichen Befragung aus den bisherigen empirischen Ergebnissen damit keine Hypothese abgeleitet werden.

3.3 Treten in einem Mixed-Mode-Design mit SAQ-Papier und SAQ-Web Mode-Effekte auf?

Ein besonderes Risiko in Mixed-Mode-Designs stellen sogenannte Mode-Effekte dar [18, 28, 32, 37, 38]. Mode-Effekte sind systematische Verzerrungen, die durch die unterschiedlichen Erhebungsformen verursacht werden. Sie entstehen durch das Befragungsmedium selbst oder durch die Befragungssituation [18, 83]. Damit ist gemeint, dass ein und dieselbe Person in Abhängigkeit von der jeweiligen Erhebungsmethode unterschiedliche Antworten auf die gleiche Frage gibt [32]. Mode-Effekten sollte mit besonderer Vorsicht begegnet werden [21], insbesondere in Längsschnittstudien. Da ein Methodenwechsel im Längsschnitt nicht hinsichtlich möglicher Methodeneffekte kontrolliert werden kann, bleibt beim Wechsel der Erhebungsmethode unklar, ob eine Veränderung innerhalb einer Person durch die Zeit oder durch den Wechsel der Erhebungsmethode bedingt ist [22, 33, 84]. Eine Möglichkeit zur Kontrolle möglicher Mode-Effekte im Längsschnitt

wäre, ein kleines Subsample in der ursprünglichen Methode zu befragen und dieses Subsample als einen Vergleich zur großen Stichprobe heranzuziehen, in der die Folgebefragung mittels anderer Methode stattgefunden hat [35]. Eine solche Herangehensweise ist aber sehr zeit- und kostenaufwändig.

Eine besondere Schwierigkeit bei der Identifizierung von Mode-Effekten ist die gleichzeitige Entstehung von Selektionseffekten [32, 85]. Mit Selektionseffekten ist gemeint, dass unterschiedliche Teile der Studienpopulation verschiedene Methoden präferieren (siehe Kapitel 3.2) oder zu einer der Erhebungsmethoden keinen Zugang haben (wie beispielsweise bei fehlendem Internet- oder Festnetzanschluss) [22, 24, 47]. In einer Vielzahl von Studien wird von kleinen, aber systematischen Unterschieden der Stichprobenzusammensetzung zwischen Erhebungsmethoden berichtet [18], weshalb zur Identifikation möglicher Mode-Effekte die Adjustierung soziodemografischer Eigenschaften beim Ergebnisvergleich sehr wichtig ist [18]. Die empirisch gefundenen Unterschiede der Stichprobeneigenschaften zwischen den beiden Erhebungsmethoden SAQ-Papier und SAQ-Web sind im Kapitel 3.2 dargestellt.

Methodenbedingte Effekte lassen sich unterteilen in [22, 33]:

- (1) Interviewereffekte, die durch die An- und Abwesenheit eines Interviewers bedingt sind und positiv im Sinne einer verbesserten Datenqualität und negativ im Sinne verzerrender Antworten sein können;
- (2) Effekte aufgrund des angesprochenen Kommunikationskanals (auditiv versus visuell) sowie
- (3) medienbezogene Effekte, wie beispielsweise der Grad an Privatheit oder Anonymität, der mit einer Erhebungsmethode assoziiert ist.

Je stärker sich die Erhebungsmethoden in ihren Eigenschaften unterscheiden, desto mehr methodenbedingte Unterschiede in den Antworten treten auf. Besonders große Unterschiede werden zwischen Interviews und Selbstausfüll-Fragebögen sowie bei Unterschieden im angesprochenen Kommunikationskanal gefunden [24]. Begünstigt werden methodenbedingte und andere verzerrende Effekte auf die Antworten von dem

sogenannten Satisficing Effekt [86, 87]. Damit ist gemeint, dass die Befragten nicht mehr die bestmöglichen Antworten geben, sondern im Sinne eines heuristischen Prozesses die minimal zufriedenstellenden Antworten, die der Survey benötigt. Dieser Effekt wird von Ermüdungserscheinungen begünstigt und kann auch im Kontext verschiedener Erhebungsmethoden auftauchen [88]. Mode-Effekte sind am größten bei sensiblen Fragen zwischen verschiedenen Modetypen, die sich durch die Präsenz einer Interviewerin beziehungsweise eines Interviewers unterscheiden. Mit sensiblen Fragen in einem Mixed-Mode-Design sollte deshalb generell vorsichtig umgegangen werden, da sich unterschiedliche Ergebnisse aufgrund eines anderen Antwortverhaltens nicht mittels Gewichtungsfaktor korrigieren lassen [89].

Im Allgemeinen gelten die beiden Selbstausfüll-Befragungsmethoden SAQ-Papier und SAQ-Web als methodenäquivalent ohne Unterschiede in den Messergebnissen [18, 29, 33, 75]. Laut einer vergleichenden Metastudie [90] weisen sie zudem die gleichen Eigenschaften des Messfehlers auf. Methodenäquivalenz ist hier wie folgt definiert: Wenn eine einzelne Nutzerin oder ein einzelner Nutzer das gleiche Instrument mit zwei Methoden ausfüllt, sind die Antworten und die daraus berechneten Scores die gleichen [91]. Begründet ist die Äquivalenz der Methoden vor allem darin, dass keine Unterschiede im genutzten Kommunikationskanal und hinsichtlich der Präsenz eines Interviewers beziehungsweise einer Interviewerin und somit der Gefahr sozial erwünschter Antworten vorhanden sind [33]. Beide Befragungsmethoden können jederzeit selbstbestimmt durchgeführt und unterbrochen werden, wodurch dem Satisficing-Effekt entgegengewirkt wird. Aufgrund der assoziierten Anonymität ist im Vergleich zu interviewbasierten Methoden eine höhere Bereitschaft vorhanden, sensible Fragen zu beantworten [22, 35, 75, 92, 93].

Im Vergleich der Online-Erhebung mit anderen Befragungsformen spielt der dabei beschriebene Effekt der „sozialen Entkontextualisierung“ eine Rolle, bei dem es aufgrund einer verminderten sozialen Präsenz zu einer reduzierten Orientierung an sozialen Normen und einer stärkeren Hervorhebung individueller Wünsche und Ziele kommt [94]. Dadurch stelle sich der Online-Modus als der „pessimistische Modus“ dar, vor

allem mit Blick auf sensible Fragen und im Vergleich zu einer interviewerbasierten Erhebung mit dem Effekt der sozialen Erwünschtheit [27, 89]. Allerdings gibt es auch den Befund häufigerer sozial erwünschter Antworten in der Online-Erhebung im Vergleich zur Papierbefragung, den die Autorinnen und Autoren im Kontext einer möglicherweise mit dem Internet assoziierten geringeren Privatheit interpretieren [95]. Im Zuge der weiteren Entwicklung sozialer Medien und der gesellschaftlichen Auseinandersetzung um Überwachungsmöglichkeiten der Internetnutzung wäre es wünschenswert, dass zukünftige Studien die These verringerter Privatheit in der SAQ-Web-Befragung versus SAQ-Papier-Befragung untersuchen.

Es gibt verschiedene Fragenformen, die nach dem bisherigen Kenntnisstand nicht durch den Erhebungsmodus beeinflusst sind. Dabei handelt es sich um Faktenfragen [33] wie soziodemografische Angaben [18] oder nicht-sensible Fragen, die ausschließlich mit ja oder nein beantwortet werden [33] sowie standardisierte Skalen im Allgemeinen [18, 37]. Eine Vielzahl empirischer Studien hat standardisierte Skalen zwischen den beiden Erhebungsmethoden SAQ-Papier und SAQ-Web verglichen, häufig vor dem Hintergrund, dass für bereits eingesetzte und standardisierte Tests in Papierform Online-Varianten entwickelt wurden. Eine Metaanalyse der seit 2007 publizierten Vergleiche zwischen SAQ-Papier und SAQ-Web von 79 standardisierten Instrumenten in insgesamt 55 Studien konstatiert inhaltlich äquivalente Ergebnisse aller Instrumente. Dabei bevorzugen die Teilnehmerinnen und Teilnehmer zu 87 % die elektronische Form [91]. Diese Ergebnisse sind allerdings nicht unmittelbar auf Surveybedingungen zu übertragen, da Unterschiede zwischen den kontrollierten Ausfüllbedingungen bei standardisierten Tests und einer klassischen Online-Befragung bestehen [31]. Aber auch populationsbasierte Surveys finden eine Äquivalenz der eingesetzten Skalen zwischen SAQ-Papier und SAQ-Web wie beispielsweise in der Methodenstudie des „Dutch Crime Victimization Survey“ [96] oder beim holländischen Arm des European Social Survey [24]. Keine Unterschiede finden sich in der psychometrischen Qualität von Rating-Skalen bei Erwachsenen [37] oder bei Jugendlichen [97] sowie in den Angaben von Extremwerten (Skalen werden an den Rändern

beantwortet) oder einem „Acquiescence-Response-Style“ (Zustimmungstendenz) zwischen den beiden Erhebungsmethoden [74].

Teilweise sind die verschiedenen Studien zu Methodeneffekten aber schwer zu generalisieren und auf die eigenen Bedarfe anzuwenden, da sie sehr stark hinsichtlich der untersuchten Population, dem Stichprobenzugang sowie den gestellten Fragen variieren [33].

Einem methodenimmanenten systematischen Messfehler bei Befragungen mittels SAQ-Papier und SAQ-Web-Modus kann mit einem sorgfältigen Design begegnet werden [36]. Die Instrumente sollten identisch umgesetzt sein, um Designeffekte zu vermeiden [18, 33, 34, 36, 50]. Es sollte ein übereinstimmendes Layout gewählt und keine Adaption an den Erhebungsmodus vorgenommen werden [71]. Dabei empfiehlt es sich, eine präferierte Methode festzulegen und das Fragebogendesign für diese Erhebungsmethode zu optimieren. Die anderen Methoden richten sich exakt daran aus und sollen so wenig wie möglich davon abweichen [35, 79]. Im Abschlussbericht der Taskforce der US Food and Drug Administration zur Testung der Messung patientenbasierter Outcomes mittels elektronischer Fragebögen und deren Implementierung in klinische Studien kommen die Autorinnen und Autoren zu dem Schluss, dass bei exakter Adaption standardisierter schriftlicher Instrumente in die elektronische Form Methodenäquivalenz zwischen SAQ-Papier und SAQ-Web erreicht werden kann [98].

Zusammenfassend lässt sich nach empirischer Befundlage sagen, dass sich Mode-Effekte in Mixed-Mode-Studien mit den Selbstausfüll-Erhebungsmodi SAQ-Papier und SAQ-Web mit einer guten Fragebogenkonstruktion vermutlich vermeiden lassen. Besondere Vorsicht und genaue Prüfung auf mögliche Mode-Effekte sollten aber sensible Fragen wegen der Verzerrungsgefahr aufgrund pessimistischerer Antworten im Online-Fragebogen erfahren.

3.4 Kann mit einem Mixed-Mode-Design SAQ-Papier und SAQ-Web die Datenqualität und Surveyökonomie erhöht werden?

Die Befragung mittels Internet gilt als sehr attraktiv für große Surveys und wird international auf-

grund ihrer – im Vergleich zu anderen Befragungsmethoden – hohen Kosteneffizienz sowie verbesserter Datenqualität [18, 40] zunehmend eingesetzt [31].

Ein zentraler Vorteil ist die unmittelbare Dateneingabe in der Online-Befragung durch die Befragten selbst. Dadurch ist die potenzielle Fehlerquelle der Dateneingabe ausgeschlossen [27, 71, 99] und es sind bereits während der Dateneingabe Prüfungen auf Pflichtangaben und Plausibilität möglich [27, 75]. Die unmittelbare Dateneingabe bietet die Möglichkeit, umgehend einen Datensatz zu schaffen und die Datenqualität zu prüfen [71, 75]. Bei der Dateneingabe eines SAQ-Papier-Fragebogens besteht dagegen eine gewisse Fehleranfälligkeit. Die Dateneingabe erfolgt entweder manuell durch geschulte Mitarbeiterinnen und Mitarbeiter oder mittels automatisierter Scanprozesse. Vor allem wenn die Befragten undeutliche Angaben machen (Freitexte, unklare Ankreuzen), können Fehler entstehen.

Ein weiterer Vorteil der Online-Befragung ist die Möglichkeit zur automatisierten Filterführung. Diese erlaubt den besseren und für die Befragten deutlich schnelleren und angenehmeren Einsatz komplexer Fragebögen durch aufwändige Filterführung [27, 69, 71]. In SAQ-Papier-Fragebögen können hingegen häufig Filterfehler entstehen, was wiederum viele unplausible Angaben und einen hohen Aufwand in der Datenbereinigung nach sich ziehen kann [75].

In der webbasierten Befragung ist es zudem möglich, die Item-Nonresponse für einzelne Fragen dadurch zu erschweren, dass die Befragten bei Nichtbeantwortung entweder gar nicht zur nächsten Frage weitergeleitet werden (Pflichtangaben) oder sie die Nichtbeantwortung der Frage zuerst bestätigen müssen, bevor die nächste Frage auf dem Bildschirm erscheint. Allerdings kann nicht verhindert werden, dass die Befragten den Gesamtfragebogen frühzeitig abbrechen und dadurch bei den hinteren Fragen im Fragebogen Item-Nonresponse entsteht [21].

In vielen empirischen Studien wurde die allgemeine Datenqualität der SAQ-Web-Befragung, wie beispielsweise vollständige oder konsistente Angaben, im Vergleich zur schriftlichen Befragung als höher beurteilt [27, 49, 67, 69, 71, 76, 99, 100]. Nur in einer vergleichenden Studie konnte kein Unterschied gefunden werden [50].

Studien zur vergleichenden Item-Nonresponse zwischen SAQ-Papier und SAQ-Web kommen zu widersprüchlichen Ergebnissen [22]. Die meisten Studien berichten eine geringere Item-Nonresponse in SAQ-Web-Surveys [z. B. 27, 45, 50, 53, 57, 67, 71, 74, 78, 97, 101, 102]. So ist ein höherer Anteil übersprungener Fragen bei sensiblen Fragen in der schriftlichen Befragung zu beobachten [103]. Die hohe Item-Nonresponse im schriftlichen Fragebogen bedeutet eine problematische Datenqualität, einen hohen Aufwand in der Datenaufbereitung und betrifft dabei auch Informationen wie Alter oder Geschlecht [75]. Manche Studien finden aber auch eine höhere Item-Nonresponse in Online-Befragungen [22, 51, 52]. Wiederum andere Studien können keinen Unterschied in der Item-Nonresponse beobachten [z. B. 37, 76, 103].

Die Kosteneffizienz webbasierter Erhebungen und/oder eine höhere Geschwindigkeit der Datenerhebung ist in einer Vielzahl von Studien bestätigt [25, 27, 29, 37, 47, 49, 50, 63, 75, 76, 99, 100, 103–105]. Es müssen keine Fragebögen verschickt werden und die hohen Kosten von SAQ-Papier-Befragungen aufgrund der nachträglichen Datenprüfung und Korrektur unplausibler Angaben entfallen [75]. In einer niederländischen Studie waren die Kosten pro Teilnehmer in der schriftlichen Befragung doppelt so hoch wie bei SAQ-Web [63]. Ein weiterer wichtiger Aspekt ist die Dauer der Datenerhebung, die im postalischen Teil einer Studie aufgrund langer Rücklaufzeiten gegenüber der Online-Erhebung deutlich länger ist [37, 76].

In Analysen zur Kosteneffizienz sollte allerdings auch die mit dem jeweiligen Erhebungsdesign erreichte Response einbezogen werden. So entstanden beispielsweise höhere Kosten pro teilnehmende Person in einer Studie in der älteren Bevölkerung aufgrund der deutlich niedrigeren Online-Teilnahme im SAQ-Web-basierten Survey als in der SAQ-Papier-Befragung [45]. In einer Response und Kosten vergleichenden Studie entstanden die höchsten Kosten im Single-Mode-Design als SAQ-Papier-Befragung und die niedrigsten Kosten im Single-Mode-Design in SAQ-Web-Form, während die Kosten des Mixed-Mode-Designs beider Erhebungsmethoden dazwischen lag [55]. Eine vergleichende Erhebung in kleinen Dörfern und ländlichen Gebieten empfiehlt unter Kostenaspekten ebenfalls die Mixed-Mode-Erhebung, in der zusätzlich zur SAQ-Papier-Befragung die Online-Erhe-

bung angeboten wird [106]. Der Vergleich zweier sequenzieller Mixed-Mode-Designs, entweder mit der SAQ-Papier- oder SAQ-Web-Befragung als erstem Modus, ergab die höheren Responsequoten in dem Design, in dem zuerst der schriftliche Fragebogen angeboten wurde. Unter Kostenaspekten kommt Börkan (2010) allerdings zu dem Schluss, dass dennoch zunächst der kostengünstige Online-Fragebogen in einem sequenziellen Erhebungsdesign angeboten werden sollte und erst im Anschluss daran der schriftliche Fragebogen [37].

Mit Blick auf die Kosteneffizienz im Rahmen eines Mixed-Mode-Designs ist allerdings zu beachten, dass die Integration eines Online-Teilnahmeangebots in eine ursprünglich ausschließlich schriftliche Befragung zunächst mit einem erhöhten Aufwand und höheren Kosten einhergeht. Es müssen die Fragebögen angepasst sowie verschiedene Alternativen im Kontakt mit den Teilnehmenden beachtet werden, ein Server eingerichtet und im Anschluss die Daten korrekt homogenisiert werden [24]. Zudem muss für die Programmierung äquivalenter Online-Fragebögen zwingend ein zeitlicher Vorlauf einkalkuliert werden [99]. Diese Kosten lohnen sich nur bei einer hinreichend großen realisierten Stichprobe mit einem hohen Anteil an Teilnehmerinnen und Teilnehmern, welche die Online-Option nutzen [27].

3.5 Zusammenfassung und Fazit

Es ist nach Studienlage nicht eindeutig zu beantworten, ob Mixed-Mode-Studiendesigns unter Einbeziehung der SAQ-Web-Befragung zusätzlich zur SAQ-Papier-Befragung unter der Perspektive des Total Survey Error attraktiv sind. Zwar konstatieren De Leeuw und Hox (2011), dass sich im Allgemeinen die Stichprobenzusammensetzung verbessert und eine geringere Nonresponse beschrieben wird. Vergleichende Studien zur Response berichten aber sowohl von verringerter, gleichbleibender und höherer Response. Eine Verbesserung der Stichprobenzusammensetzung wird in erster Linie bezüglich der Verteilung nach Alter und Geschlecht beschrieben. Es kommt allerdings sehr darauf an, in welcher Population die Methoden eingesetzt und gemischt werden. Die Messergebnisse beider Erhebungsmethoden werden

im Allgemeinen als äquivalent bewertet [29]. Sehr wichtig ist die identische Umsetzung der Fragebögen, um Designeffekte zu vermeiden [18, 33, 34, 36]. Sollten modespezifische Optimierungen oder andere Umsetzungen von Fragen erfolgen, sollte die Äquivalenz der Ergebnisse unbedingt getestet [29] und nach Möglichkeit adjustiert [33] werden. Der eindeutigste Vorteil liegt – aber nur in einem sequenziellen Mixed-Mode-Design – im Bereich der Surveyökonomie und der besseren Datenqualität durch das zusätzliche Angebot eines Online-Fragebogens [37]. Dies gilt allerdings nur für eine große realisierte Stichprobe und einen hinreichend hohen Anteil von Online-Teilnehmenden.

4 Die methodische Pilotstudie: Einbettung, Design und Inhalte

4.1 Einbindung der Pilotstudie in den Pretest für KiGGS Welle 2

Der Pretest für KiGGS Welle 2 wurde systematisch in einen Pretest zur Untersuchung einerseits und die methodische Pilotstudie zur Befragung andererseits unterteilt. Beide Vorhaben beschränkten sich auf die Zielpopulation von 0 bis 17 Jahren. Die Pilotstudie zur Befragung wurde von August bis Dezember 2013 durchgeführt.

4.2 Zielstellungen der methodischen Pilotstudie

4.2.1 Vergleich unterschiedlicher Erhebungsdesigns

In der methodischen Pilotstudie wurden in verschiedenen Erhebungsdesigns unterschiedliche Abfolgen von Teilnahmeangeboten realisiert, um deren Effektivität mit Blick auf die Stichprobenausschöpfung und Repräsentativität zu betrachten. Insbesondere sollte geprüft werden, ob in Abhängigkeit der verschiedenen Erhebungsdesigns das zur schriftlichen Befragung alternative Angebot der Online-Befragung Personen zur Teilnahme bewegen kann, die sonst möglicherweise nicht teilgenommen hätten. Relevant waren dabei auch Kosten-Nutzen-Überlegungen im Sinne einer Nutzung eines der beiden Erhebungsmodi durch die Teilnehmerinnen und Teilnehmer.

4.2.2 Vergleich unterschiedlicher Erhebungsmodi

Um systematische Fehlerquellen in der Mixed-Mode-Erhebung bestmöglich ausschließen zu können, wurde in der methodischen Pilotstudie neben dem Vergleich der Erhebungsdesigns auch die Stichprobenzusammensetzung innerhalb der Erhebungsmodi (über die Erhebungsdesigns hinweg) betrachtet. Außerdem wurde nach diesem Schema auch das Auftreten von Item-Nonresponse als Indikator für Datenqualität sowie die inhaltliche Vergleichbarkeit der Ergebnisse zentraler Gesundheitsindikatoren (Prävalenzschätzungen und Mittelwerte) geprüft.

4.3 Studiendesign der methodischen Pilotstudie

4.3.1 Stichprobe

Die Bruttostichprobe der methodischen Pilotstudie umfasste 11.140 Adressen von Kindern und Jugendlichen im Alter von 0 bis 17 Jahren. Gezogen wurden die Adressen in jeweils fünf Orten (Sample Points) aus den Bundesländern Berlin, Brandenburg, Sachsen, Schleswig-Holstein und Baden-Württemberg. Die Sample Points wurden in Bezug auf Region und Gemeindegröße so ausgewählt, dass sowohl städtische als auch ländliche Gemeinden in den neuen und alten Bundesländern vertreten waren. Die Adressen der potenziellen Teilnehmerinnen und Teilnehmer wurden per Zufallsverfahren auf die verschiedenen Erhebungsdesigns aufgeteilt (Abbildung 3).

4.3.2 Kontrollgruppendesign

Abbildung 3 stellt das Kontrollgruppendesign mit seinen verschiedenen Angebotsverfahren zur Teilnahme dar. Die Kinder und Jugendlichen wurden ab 11 Jahren selbst befragt und erhielten grundsätzlich die gleichen Teilnahmeangebote wie ihre Eltern, entsprechend der Zuweisung zu einem der Erhebungsdesigns. Die Gewinnung der Teilnehmerinnen und Teilnehmer begann in allen Erhebungsdesigns mit einer postalischen Einladung zur Studienteilnahme. Daraufhin wurde ein Erinnerungsschreiben versendet, wenn circa zweieinhalb Wochen nach dem Versand der Einladung

noch keine Rückmeldung erfolgt war. Erfolgte hierauf ebenfalls keine Reaktion, wurde versucht, eine Telefonnummer der betreffenden Person zu recherchieren. Bei erfolgreicher Recherche wurde circa vier Wochen nach dem Versand des Einladungsschreibens bis zu fünf Mal versucht, telefonisch Kontakt herzustellen.

Den Einladungsschreiben lag eine schriftliche Einwilligungserklärung bei, die in einem beigelegten Rückumschlag portofrei ans RKI zurückgesendet werden sollte. Eine Ausnahme bildet die Einladung an die Eltern von 0- bis 10-jährigen Kindern im sequenziellen Erhebungsdesign. Ab einem Alter von 14 Jahren wurden neben den Eltern auch die Jugendlichen selbst um ihre Einwilligung gebeten. In der Altersgruppe der 11- bis 17-Jährigen war stets eine schriftliche Einwilligung erforderlich – auch im Falle einer Teilnahme an der Online-Befragung (durch die Eltern und/oder Kinder beziehungsweise Jugendlichen). Im Falle einer Wahlmöglichkeit zwischen den verschiedenen Teilnahmeangeboten konnte die Entscheidung auf der schriftlichen Einwilligungserklärung angegeben werden. Die Fragebögen für die Kinder und Jugendlichen wurden erst zugänglich gemacht, wenn die erforderlichen Einwilligungen vorlagen. Voraussetzung für das Erfüllen der Falldefinition (siehe Kapitel 5.1.1) war das Vorliegen aller erforderlichen Einwilligungserklärungen.

An alle Personen, die letztlich weder auf dem postalischen Weg noch über ein Telefongespräch zur Teilnahme gewonnen werden konnten, wurde abschließend ungefähr zehn Wochen nach der Einladung ein zweiseitiger Nichtteilnehmerfragebogen (Nonresponder-Fragebogen) verschickt. Dieser beinhaltete vier Fragen zur Gesundheit, sechs Fragen zu Survey-Einstellungen sowie verschiedene Fragen zur Soziodemografie.

Zur Steigerung der Teilnahmemotivation konnten die Teilnehmerinnen und Teilnehmer auf in der Einladung und Erinnerung mitgeschickten Gutscheinformularen beziehungsweise online wählen, von welchem dreier Anbieter (Karstadt, MediaMarkt und H&M) sie gerne einen Einkaufsgutschein im Wert von 10 Euro pro ausgefülltem Gesundheitsfragebogen erhalten würden (Eltern und Kinder/Jugendliche wählten getrennt). Nach Eingang des Gesundheitsfragebogens wurde der Gutschein zeitnah versendet. Dabei wurde an ggf. fehlende Unterlagen (z. B. Ernährungsfragebogen,

elterliche Einwilligungserklärung bei 0- bis 10-jährigen Kindern) erinnert.

Die beschriebenen Einladungsprozesse wurden über Datenbankanwendungen durchgeführt, über die auch die Registrierung von eingehenden Mitteilungen und Unterlagen vorgenommen werden konnte. Die Testung der Datenbankanwendungen ist im Rahmen der surveyökonomischen Betrachtungen (z. B. Praktikabilität der verschiedenen Angebotsverfahren) eine wichtige Komponente der Pilotstudie.

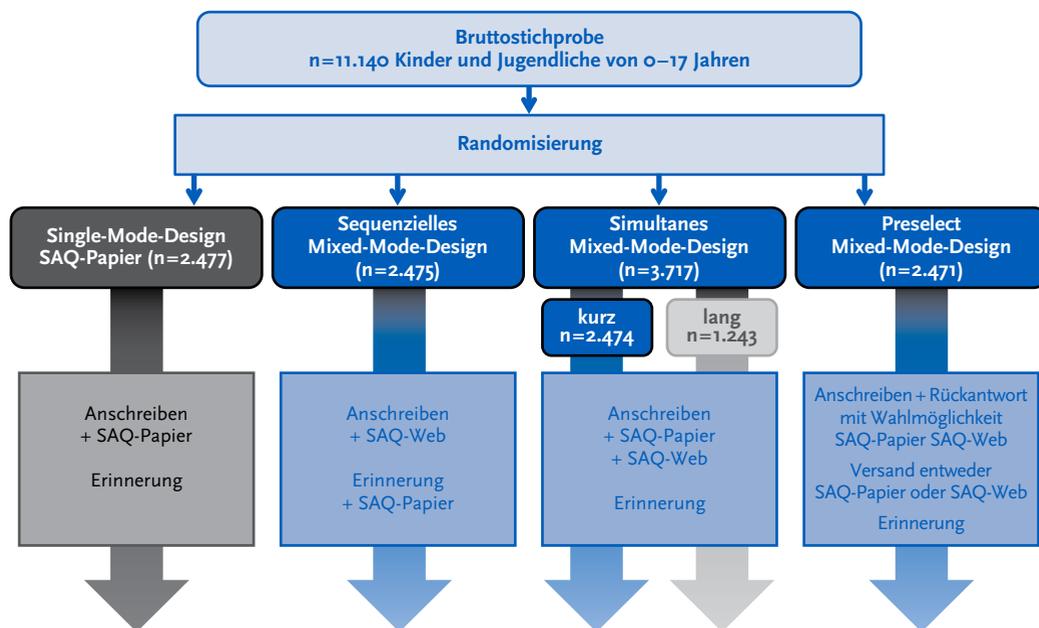
- (1) Single-Mode (Kontrollgruppe) – ausschließliches Angebot schriftlicher Fragebögen in Papierform, die direkt mit dem Einladungsschreiben zugesendet wurden.
- (2) Mixed-Mode Sequenziell – sequenzielles Teilnahmeangebot, bei dem im Einladungsschreiben zunächst nur die Zugangsdaten zur Online-Befragung (Webseite, persönlicher Zugangscode) enthalten waren; reagierten die Eingeladenen darauf nicht, wurde mit der Erin-

nerung zusätzlich die Papierversion der Fragebögen zugesendet.

- (3) Mixed-Mode Simultan – simultanes Teilnahmeangebot, bei dem mit der Einladung die schriftlichen Fragebögen in Papierform und zugleich die Zugangsdaten zur Online-Befragung zugesendet wurden.

Die in Abbildung 3 dargestellte Trennung in die Gruppen „kurz“ und „lang“ zielt darauf ab, den Einfluss verschiedener Längen des Gesundheitsfragebogens auf die Response zu überprüfen. Der kürzere Fragebogen kam dabei auch in allen anderen Erhebungsdesigns zum Einsatz, wodurch nur die Gruppe „Mixed-Mode Simultan kurz“ hinsichtlich der Stichprobenzusammensetzung mit den anderen Erhebungsdesigns vergleichbar ist. Sofern nicht explizit anders benannt, wird in diesem Bericht von der Gruppe „Mixed-Mode Simultan kurz“ berichtet. Für die Vergleiche der Erhebungsmodi können die Teilnehmerinnen und Teilnehmer der Gruppe „Mixed-Mode Simultan

Abbildung 3
Kontrollgruppendesign der methodischen Pilotstudie



lang“ jedoch herangezogen werden. Angaben zur Gesamtstudie wie Fallzahlen und Responsequote schließen die Gruppe „Mixed-Mode Simultan lang“ ein.

4. Mixed-Mode Preselect – Teilnahmeangebot mit Vorauswahl, bei dem die Eingeladenen sowohl im Einladungs- als auch Erinnerungsschreiben gebeten wurden, auf einem portofrei zurücksendenden Antwortformular ihre gewünschte Befragungsform zu wählen, woraufhin ihnen die Zugangsdaten zur Online-Befragung oder die Papierversion der Fragebögen zugesendet wurden.

4.4 Studieninhalte

Die methodische Pilotstudie war Teil des allgemeinen Pretests von KiGGS Welle 2. Es wurden alle für diese Erhebungswelle geplanten Themenkomplexe in die Fragebögen aufgenommen (Abbildung 4).

Es wurden verschiedene soziodemografische Merkmale wie Alter, Geschlecht, Familienform/Familienstand, schulische und berufliche Ausbildung, Erwerbsstatus und berufliche Stellung von den Teilnehmenden und Eltern, Haushaltsgröße und -zusammensetzung, Haushaltsnettoeinkommen der Familie und Migrationshintergrund erhoben. Mit Blick auf die Lebenswelten der Kinder und Jugendlichen wurde zum einen die außerfamiliäre Betreuung erfragt, zum anderen wurden Umweltfaktoren erhoben wie Wohnraum, Spiel- und Sportmöglichkeiten oder Lärmbelastung [16].

Aus der Vielzahl abgefragter Gesundheitsindikatoren (Abbildung 4) werden im Folgenden nur diejenigen detailliert beschrieben, die hinsichtlich möglicher Mode-Effekte analysiert worden sind. Zentrales Kriterium für die Auswahl ist deren Public-Health-Relevanz für die bevölkerungsbezogene Berichterstattung zur Kinder- und Jugendgesundheit.

4.4.1 Allgemeiner Gesundheitszustand von Kindern und Jugendlichen

Die subjektive Einschätzung des allgemeinen Gesundheitszustands wurde über die Elternein-

schätzung für Kinder und Jugendliche im Alter von 3 bis 17 Jahren sowie über die Selbsteinschätzung der Kinder und Jugendlichen im Alter von 11 bis 17 Jahren entsprechend der von der Weltgesundheitsorganisation (WHO) empfohlenen Formulierung [107] erfragt: „Wie würden Sie den Gesundheitszustand Ihres Kindes im Allgemeinen beschreiben?“ beziehungsweise „Wie würdest du deinen Gesundheitszustand im Allgemeinen beschreiben?“. Die fünfstufig angelegte Antwortskala wurde für die statistischen Analysen dichotomisiert in „sehr gut/gut“ und „mittelmäßig/schlecht/sehr schlecht“ [108, 109].

Chronische Krankheiten bei 0- bis 17-jährigen Kindern und Jugendlichen wurden neben der Abfrage von Prävalenzen ärztlich diagnostizierter Krankheiten mittels Elternangabe mit der für Kinder adaptierten Frage aus dem Minimum European Health Instrument (MEHM) operationalisiert „Hat Ihr Kind eine oder mehrere lang andauernde, chronische Krankheiten oder Gesundheitsprobleme? (Chronische Krankheiten sind lang andauernde Erkrankungen, die regelmäßiger Behandlung und Kontrolle bedürfen, zum Beispiel Asthma, Epilepsie, Diabetes oder Herz-erkrankungen)“ [110].

Gesundheitliche Beeinträchtigungen bei 0- bis 17-jährigen Kindern und Jugendlichen wurden ebenfalls mittels Elternangabe erhoben. Hierfür wurde die Frage aus dem Kurzfragebogen Children with Special Health Care Needs (CSHCN-Screener) [111] gestellt, ob das Kind „aus gesundheitlichen Gründen in irgendeiner Art und Weise eingeschränkt oder daran gehindert ist, Dinge zu tun, die die meisten gleichaltrigen Kinder tun können?“ (Antwortkategorien: „ja“, „nein“) [112].

4.4.2 Wichtige Erkrankungen bei Kindern und Jugendlichen

Das Vorliegen einer ADHS-Diagnose wurde von den Eltern für Kinder und Jugendliche im Alter von 3 bis 17 Jahren erfasst. Die Frage lautete: „Wurde bei Ihrem Kind jemals eine Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung, auch als ADHS bzw. ADS bezeichnet, festgestellt?“ [113, 114].

In Bezug auf allergische Erkrankungen wurde die Lebenszeitprävalenz von Asthma bronchiale, Heuschnupfen sowie Neurodermitis bei Kindern

Abbildung 4
Studieninhalte der methodischen Pilotstudie

<p>Körperliche Gesundheit</p> <ul style="list-style-type: none"> ▶ Subjektive Gesundheit ▶ Chronische Krankheiten ▶ Unfälle und Verletzungen ▶ Schmerzen ▶ Infektionskrankheiten ▶ Allergien/Asthma bronchiale ▶ Herzerkrankung ▶ Epilepsie ▶ Diabetes ▶ Behinderung/Einschränkungen ▶ Körpergröße und Gewicht ▶ Reife, Entwicklung 	<p>Psychische Gesundheit</p> <ul style="list-style-type: none"> ▶ Psychische Auffälligkeiten ▶ Psychische Störungen ▶ ADHS ▶ Gesundheitsbezogene Lebensqualität ▶ Essstörungen ▶ Subjektive Körperwahrnehmung ▶ Familienklima ▶ Personale Ressourcen ▶ Selbstwirksamkeit 	<p>Gesundheitsverhalten</p> <ul style="list-style-type: none"> ▶ Ernährung ▶ Körperliche Aktivität ▶ Sport ▶ Mediennutzung ▶ Alkoholkonsum ▶ Tabakkonsum ▶ Passivrauchbelastung ▶ Schlafen ▶ Geschlechtsverkehr
<p>Elterliche Einflussfaktoren</p> <ul style="list-style-type: none"> ▶ Sport ▶ Tabakkonsum ▶ Rauchen in Schwangerschaft/Stillzeit ▶ Stillverhalten ▶ Körpergröße und Gewicht ▶ Allergien ▶ Psychische Erkrankung ▶ Elterliche Belastung 	<p>Soziale Determinanten</p> <ul style="list-style-type: none"> ▶ Subjektiver Sozialstatus ▶ Migrationshintergrund ▶ SES Elternhaus ▶ Bildung des Kindes/Jugendlichen ▶ Familiärer Hintergrund ▶ Umwelteinflüsse (Lärm, Wohnen) ▶ Krankenversicherung ▶ Vorschulische Betreuung ▶ Soziale Unterstützung 	<p>Prävention/Versorgung</p> <ul style="list-style-type: none"> ▶ Inanspruchnahme: Ärztinnen und Ärzte, Therapeutinnen und Therapeuten ▶ Kieferorthopädische Behandlung ▶ Zahnpflege ▶ Krankenhausaufenthalte ▶ U-Untersuchungen ▶ Impfen ▶ Health Literacy ▶ Frühe Hilfen ▶ Operationen

ADHS = Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung; SES = Sozioökonomischer Status

und Jugendlichen im Alter von 0 bis 17 Jahren durch die Eltern erfasst. Die Eltern wurden dabei gefragt, ob die Krankheit bei ihrem Kind jemals ärztlich festgestellt wurde [115, 116].

Wiederholt auftretende Schmerzen wurden bei Kindern und Jugendlichen im Alter von 11 bis 17 Jahren selbst erfasst. Dabei wurde gefragt: „Hatest du in den letzten 3 Monaten eine der folgenden Schmerzarten?“ – Kopfschmerzen, Rückenschmerzen, Ohrenscherzen, Augenschmerzen, Bauchschmerzen, Regelschmerzen (bei Mädchen), Schmerzen im Unterleib, Armschmerzen/Handschmerzen, Beinschmerzen/Fußschmerzen, Schmerzen im Brustkorb, Halsschmerzen, Zahnschmerzen, andere Schmerzen (offenes Antwort-

feld). Die drei dazugehörigen Antwortkategorien wurden für die statistischen Analysen von Rücken- und Kopfschmerzen sowie von „irgendwelche Schmerzen“ dichotomisiert in „ja, wiederholt“ vs. „ja, einmalig/nein“ [117].

4.4.3 Psychische Auffälligkeiten und Beeinträchtigungen bei Kindern und Jugendlichen

Zur Erhebung psychischer Auffälligkeiten wurde der international anerkannte Strengths and Difficulties Questionnaire (SDQ) eingesetzt, der aus Symptom- und Impactfragebogen besteht [118, 119].

Für den Altersbereich von 3 bis 17 Jahre liegen Elternangaben vor, Kinder und Jugendliche im Alter von 11 bis 17 Jahren haben den Fragenkatalog zu psychischen Auffälligkeiten selbst ausgefüllt.

Im Symptomfragebogen werden 25 Items abgefragt, die sich auf fünf Subskalen beziehen, davon die vier Problemskalen „emotionale Probleme“, „Verhaltensprobleme“, „Hyperaktivitätsprobleme“ und „Probleme im Umgang mit Gleichaltrigen“ sowie die Stärkenskala „prosoziales Verhalten“. Für alle fünf SDQ-Skalen kann ein Testwert gebildet und aus der Summe der vier Problemskalen ein Gesamtproblemwert ermittelt werden. Kinder und Jugendliche, die sich im Gesamtproblemwert anhand der Cut-off-Werte einer deutschen Normstichprobe [120, 121] als grenzwertig auffällig oder auffällig klassifizieren lassen, werden als SDQ Symptome Risikogruppe zusammengefasst [122, 123].

Das Impactsupplement beginnt mit der Eingangsfrage, ob das Kind keine, leichte, deutliche oder massive Schwierigkeiten in den vier Bereichen Stimmung, Verhalten, Konzentration und Umgang mit anderen habe. Werden mindestens leichte Schwierigkeiten angegeben, schließen sich Zusatzfragen zum subjektiven Leidensdruck des Kindes sowie zu subjektiv wahrgenommenen sozialen Beeinträchtigungen in vier alltäglichen Lebensbereichen an, aus denen mittels Addition der sogenannte Impactscore gebildet wird [119]. Dieser kann ebenfalls in auffällig oder grenzwertig klassifiziert werden, beide Gruppen werden als SDQ Impact Risikogruppe zusammengefasst. Bei der Risikogruppe liegen in mindestens zwei Lebensbereichen deutliche oder in mindestens einem Lebensbereich schwere Beeinträchtigungen (auffällig) oder in mindestens einem Lebensbereich eine deutliche Beeinträchtigung (grenzwertig) des Kindes beziehungsweise Jugendlichen aufgrund psychischer Probleme vor [123].

4.4.4 Gesundheitsbezogene Lebensqualität von Kindern und Jugendlichen

Zur Erfassung der gesundheitsbezogenen Lebensqualität wurden in Abhängigkeit vom Alter zwei verschieden lange Formen des KIDSCREEN-Fragebogens eingesetzt [124, 125]. Die KIDSCREEN-Fragebögen erfassen mit einer verschieden großen

Anzahl von Items Dimensionen gesundheitsbezogener Lebensqualität von Kindern und Jugendlichen entweder aus Eltern- oder Selbstsicht.

Für Kinder im Alter von 7 bis 10 Jahren beantworteten die Eltern den KIDSCREEN-10-Fragebogen. Dieser stellt einen globalen Lebensqualitätsscore dar, der das Wohlbefinden in den letzten 7 Tagen anhand von 10 Einzelfragen zu fünf Dimensionen des Wohlbefindens – Körper, Psyche, Eltern, Freunde und Schule – erfragt. Bei den 11- bis 17-Jährigen wurde hingegen der KIDSCREEN-27-Fragebogen im Selbsturteil eingesetzt, der das Wohlbefinden in den letzten sieben Tagen mit 27 Einzelfragen zu den fünf genannten Dimensionen erfasst, die gesondert ausgewiesen werden. Die Antwortkategorien geben die Häufigkeit beziehungsweise Intensität an. Die Angaben zu den jeweiligen Antworten werden aufsummiert und in t-Werte (Mittelwert 50 und Standardabweichung 10) transformiert, wobei höhere Werte auf eine bessere gesundheitsbezogene Lebensqualität schließen lassen [126]. Die Berechnung der Scores erfolgt mit dem vorgegebenen Auswertungsprogramm, in dem fehlende Werte, wenn möglich, ersetzt werden.

Die Kinder und Jugendlichen können anhand der berechneten t-Werte in 3 Gruppen eingeteilt werden: Eine sehr gute Lebensqualität haben diejenigen, deren Werte um mehr als eine Standardabweichung höher als der Mittelwert liegen (> 60), eine gute Lebensqualität weisen jene mit Mittelwerten zwischen 40 und 60 auf und eine geringe Lebensqualität haben diejenigen, deren Werte um mehr als eine Standardabweichung niedriger als der Mittelwert liegen (< 40) [127, 128].

4.4.5 Adipositas bei Kindern, Jugendlichen und deren Eltern

Zur Definition von Übergewicht und Adipositas im Kindes- und Jugendalter wird der Body Mass Index (BMI: Körpergewicht in kg/Körpergröße in m²) eingesetzt [129]. Körpergröße und -gewicht wurden von Kindern im Alter von 0 bis 10 Jahren durch die Eltern und im Alter von 11 bis 17 Jahren von den Kindern und Jugendlichen selbst im Fragebogen in Zentimetern respektive Kilogramm angegeben. Da im Kindes- und Jugendalter keine festen Grenzwerte zur Definition von Übergewicht

und Adipositas vorliegen, ist es üblich, alters- und geschlechtsabhängige Perzentile einer Referenzpopulation als Grundlage der Definition von Übergewicht und Adipositas zu verwenden [130]. Die statistischen Analysen beruhen auf den Referenzwerten nach Kromeyer-Hauschild et al. (2001), die von der Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter [131] in Deutschland zur Nutzung empfohlen werden. Danach sind Mädchen und Jungen mit einem BMI-Wert oberhalb des 90. alters- und geschlechtsspezifischen Perzentils übergewichtig, diejenigen mit einem BMI-Wert oberhalb des 97. Perzentils adipös [132, 133].

Um Aussagen über die Verbreitung von Übergewicht und Adipositas bei den Eltern treffen zu können, wurden die Eltern nach Körpergröße und -gewicht gefragt, um auf dieser Basis den BMI berechnen zu können. Bei Erwachsenen erfolgt – anders als bei Kindern und Jugendlichen – die Definition von Übergewicht und Adipositas anhand fester Grenzwerte. Nach der WHO ist Übergewicht definiert als $\text{BMI} \geq 25 \text{ kg/m}^2$ und Adipositas als $\text{BMI} \geq 30 \text{ kg/m}^2$ [134].

Da in den Online-Fragebögen Plausibilitätschecks in den Gewichts- und Größenangaben hinterlegt sind, wurden davon abweichende Angaben in den Papierfragebögen ebenfalls als nicht plausible Angaben gewertet und als fehlende Werte (Missings) kodiert.

4.4.6 Tabakkonsum von Kindern, Jugendlichen und deren Eltern

Das Rauchverhalten von Kindern und Jugendlichen im Alter von 11 bis 17 Jahren wurde mit der Frage „Rauchst du zurzeit?“ im Selbsturteil erhoben. Die dazugehörige Antwortskala war fünfstufig angelegt: „nein“, „täglich“, „mehrmals pro Woche“, „einmal pro Woche“ und „seltener“ [135, 136]. Unter aktuellem Rauchen der Jugendlichen wird jeglicher, das heißt auch gelegentlicher Tabakkonsum verstanden.

Zur Erfassung der 12-Monatsprävalenz von Wasserpfeifenkonsum wurden alle 11- bis 17-jährigen Kinder und Jugendlichen zunächst gefragt: „Hast du schon einmal eine Wasserpfeife oder Shisha geraucht?“ (Antwortkategorien: „ja“, „nein“). Jugendliche, die diese Frage bejaht haben, wurde als nächstes die Frage gestellt, ob sie auch in den

letzten 12 Monaten Wasserpfeife oder Shisha geraucht haben (Antwortkategorien: „ja“, „nein“) [137].

Das aktuelle Rauchen des befragten Elternteils wurde über die Frage erhoben, ob sie/er zurzeit raucht. Die Antwortmöglichkeiten lauteten: „ja, täglich“, „ja, gelegentlich“ und „nein“ und wurden ebenfalls dichotomisiert in Rauchen ja/nein – unabhängig von der Regelmäßigkeit – ausgewertet. Darüber hinaus wurde das Rauchverhalten der Partnerin/ des Partners erfragt.

Mit Blick auf den elterlichen Tabakkonsum wurde zudem das mütterliche Rauchen während der Schwangerschaft und Stillzeit mit zwei Fragen erhoben. Zunächst wurde gefragt, ob die Mutter des Kindes während der Schwangerschaft geraucht hat. Die Antwortskala war dreistufig: „ja, regelmäßig“, „ja, ab und zu“ und „nein, nie“. Im Anschluss daran wurde gefragt, ob die Mutter des Kindes während der Stillzeit geraucht hat. Die Antwortskala war vierstufig: „ja, regelmäßig“, „ja, ab und zu“, „nein, nie“ und „hat nicht gestillt“. Die Antworten wurden jeweils dichotomisiert in diejenigen, die während der Schwangerschaft beziehungsweise Stillzeit geraucht (regelmäßig/ ab und zu) haben und diejenigen, die nicht geraucht haben [138].

4.4.7 Alkoholkonsum von Kindern und Jugendlichen

Der Alkoholkonsum von Kindern und Jugendlichen im Alter von 11 bis 17 Jahren wurde anhand von vier Fragen erhoben.

Um zunächst die Lebenszeitprävalenz des Alkoholkonsums zu erfassen, wurden alle Jugendlichen gefragt: „Hast du schon einmal Alkohol getrunken?“ (Antwortkategorien: „ja“, „nein“).

Danach wurde das aus drei Fragen bestehende, international etablierte AUDIT-C-Instrument eingesetzt [139]. Hierbei handelt es sich um einen Kurzfragebogen des Alcohol Use Disorders Identification Test (AUDIT), der zur Erfassung des riskanten Alkoholkonsums und zur Verbreitung des Rauschtrinkens verwendet wird. Die drei Fragen des AUDIT-C-Instruments sind: 1. „Wie oft nimmst du ein alkoholisches Getränk zu dir?“ (Antwortkategorien: „nie“, „einmal pro Monat oder seltener“, „2 bis 4 mal im Monat“, „2 bis 3 mal

pro Woche“ und „4 mal pro Woche oder öfter“), 2. „Wenn du Alkohol trinkst, wie viele alkoholische Getränke trinkst du dann üblicherweise an einem Tag?“ (Antwortkategorien: „1 bis 2“, „3 bis 4“, „5 bis 6“, „7 bis 9“ und „10 oder mehr alkoholische Getränke“) und 3. „Wie oft trinkst du sechs oder mehr alkoholische Getränke bei einer Gelegenheit, z. B. auf einer Party?“ (Antwortkategorien: „nie“, „seltener als einmal im Monat“, „jeden Monat“, „jede Woche“ und „jeden Tag oder fast jeden Tag“). Entsprechend den Instrumentenvorgaben wurden die Antwortkategorien dieser drei Einzelfragen jeweils aufsteigend mit Punktwerten von 0 bis 4 versehen und anschließend aufsummiert [139, 140]. Die Spannweite des auf diese Weise gebildeten Summenscores reicht demnach von 0 bis 12. Von riskantem Alkoholkonsum ist bei einem Summenscore von ≥ 4 bei Mädchen und ≥ 5 bei Jungen auszugehen [141, 142]. Auf Basis der dritten Einzelfrage kann zudem die Verbreitung des Rauschtrinkens bestimmt werden. Von regelmäßigem Rauschtrinken ist die Rede, wenn die Jugendlichen angaben, mindestens einmal im Monat oder häufiger sechs oder mehr alkoholische Getränke bei einer Gelegenheit zu trinken [135, 136].

4.4.8 Körperliche Aktivität und Sport bei Kindern, Jugendlichen und deren Eltern

Fragen zum Ausmaß körperlicher Aktivität wurden bei Kindern und Jugendlichen im Alter von 3 bis 17 Jahren gestellt. Für Kinder im Alter von 3 bis 10 Jahren füllte ein Elternteil die Fragen zur körperlichen Aktivität aus, Kinder und Jugendliche ab dem elften Lebensjahr beantworteten die Fragen selbst. Da die WHO [143] für Kinder und Jugendliche eine tägliche körperliche Aktivität von mindestens 60 Minuten bei mäßiger bis anstrengender Intensität empfiehlt (zusätzlich sollten Übungen zur Kräftigung der Muskulatur durchgeführt werden), wurden die Eltern und Jugendlichen im Rahmen des auszufüllenden Fragebogens gefragt: „An wie vielen Tagen einer normalen Woche ist Ihr Kind/bist du für mindestens 60 Minuten am Tag körperlich aktiv?“. Die acht Antwortkategorien reichten von „an keinem Tag“ bis zu „7 Tage“.

In den Auswertungen nach Erhebungsmodus wurden die Anteile an Kindern und Jugendlichen gegenübergestellt, die täglich mindestens 60

Minuten körperlich aktiv sind und damit körperlich aktiv analog der WHO-Empfehlung.

Zum anderen wurden die Anteile an Kindern und Jugendlichen mit relativ geringer körperlicher Aktivität verglichen. Hierunter werden all jene Mädchen und Jungen gefasst, die an weniger als zwei Tagen pro Woche für mindestens 60 Minuten pro Tag körperlich aktiv sind [144, 145].

Sportliche Aktivität wurde ebenfalls bei Kindern und Jugendlichen im Alter von 3 bis 17 Jahren erhoben. Für Kinder im Alter von 3 bis 10 Jahren beantwortete erneut ein Elternteil die Fragen, während Kinder und Jugendliche im Alter von 11 bis 17 Jahren die Angaben selbst machten. Die Frage, die im Rahmen des auszufüllenden Fragebogens gestellt wurde, lautete: „Treibt Ihr Kind/treibst du Sport?“ (Antwortkategorien: „ja“, „nein“). Sport in Kindertageseinrichtungen und Schulsport sollte dabei nicht berücksichtigt werden [145].

Ob die Eltern in ihrem Alltag sportlich aktiv sind, wurde im Rahmen des auszufüllenden Fragebogens mit der Frage erhoben: „Wie oft treiben Sie selbst Sport? (Denken Sie dabei bitte an die letzten 3 Monate)“. Die dazugehörigen Antwortkategorien lauteten: „keine sportliche Betätigung“, „weniger als eine Stunde in der Woche“, „regelmäßig 1–2 Stunden in der Woche“, „regelmäßig 2–4 Stunden in der Woche“ und „regelmäßig mehr als vier Stunden in der Woche“. Die Antworten wurden dichotomisiert in regelmäßig Sport ja (d. h. regelmäßig mindestens 2–4 Stunden Sport pro Woche) und nein (das heißt keinen Sport oder maximal 1–2 Stunden Sport pro Woche).

4.4.9 Mediennutzung von Kindern und Jugendlichen

Informationen zum Mediennutzungsverhalten von 3- bis 10-jährigen Kindern wurden über die Eltern erhoben, 11- bis 17-jährige Kinder und Jugendliche haben hierüber selbst Auskunft gegeben. Die Eltern und Jugendlichen wurden hierbei gefragt: „Wie lange beschäftigt sich Ihr Kind/beschäftigt du dich durchschnittlich pro Tag mit folgenden Dingen?“ – „Fernsehen und Video/DVD“, „Computer/Internet“, „Spielkonsole oder Computerspiele“, „Sonstige Computernutzung und Internet“ sowie „Bücher lesen“. Die vorgegebenen Antwortkategorien lauteten: „gar

nicht“, „bis zu 1 Stunde“, „bis zu 2 Stunden“, „bis zu 3 Stunden“, „bis zu 4 Stunden“ und „mehr als 4 Stunden“. Die Kinder und Jugendlichen selbst wurden zudem nach der Nutzung von sozialen Medien pro Tag gefragt. Es werden im Rahmen der statistischen Analysen die Anteile der Kinder und Jugendlichen zwischen den Erhebungsmodi verglichen, die jeden Tag mehr als zwei Stunden mit den genannten Medien verbringen [145, 146].

4.4.10 Inanspruchnahme ärztlicher Leistungen durch Kinder und Jugendliche

Die Inanspruchnahme von niedergelassenen Ärztinnen und Ärzten in den letzten 12 Monaten wurde für 0- bis 13-jährige Kinder durch die Eltern erfasst, Jugendliche ab 14 Jahren gaben hierüber selbst Auskunft. Es wurde eine Liste verschiedener Ärztinnen und Ärzte (Kinderheilkunde, Allgemeinmedizin, Innere Medizin, Augen- und Hals-Nasen-Ohren-Heilkunde, Orthopädie, Gynäkologie, Neurologie, Psychiatrie, Chirurgie, Dermatologie, Radiologie, Urologie und Zahnärztin oder Zahnarzt/ Kieferorthopädie) präsentiert, bei der angekreuzt werden konnte, ob und wie häufig die jeweilige Arztgruppe in den letzten 12 Monaten in Anspruch genommen wurde.

Analysiert werden die Anteile an Kindern und Jugendlichen, die mindestens einmal eine Kinderärztin oder einen Kinderarzt, zahnärztliche beziehungsweise kieferorthopädische Leistungen oder mindestens eine der genannten Facharztgruppen in den letzten 12 Monaten in Anspruch genommen haben (Inanspruchnahmequote) [147, 148].

Weiterhin wurden die 11- bis 17-Jährigen selbst gefragt, ob sie sich aktuell in kieferorthopädischer Behandlung befinden (Antwortkategorien: „ja“, „nein“).

4.5 Methodenübergreifende Operationalisierung der Befragungsunterlagen

Im Rahmen der methodischen Pilotstudie zu KiGGS Welle 2 kamen insgesamt acht verschiedene Fragebögen zum Einsatz, die entweder von den Eltern oder von den Kindern und Jugendlichen ab dem elften Lebensjahr selbst ausgefüllt wurden (Tabelle 1). Die Befragung der Eltern

umfasste dabei eine Befragung zum Gesundheitszustand des Kindes (Gesundheitsfragebogen, verfügbar für mehrere Altersgruppen), zu ärztlich diagnostizierten Krankheiten des Kindes (Krankheitsfragebogen, kurze oder lange Fassung) (vgl. Kapitel 4.3.2) und zum kindlichen Ernährungsverhalten (Ernährungsfragebogen, Food Frequency Questionnaire (FFQ), verfügbar für zwei Altersgruppen). Bei den Kindern und Jugendlichen im Alter von 11 bis 17 Jahren wurde ein Fragebogen zum Gesundheitszustand (Gesundheitsfragebogen) und zum Ernährungsverhalten (Ernährungsfragebogen) eingesetzt.

Im Zuge der Fragebogenentwicklung für die schriftliche und die webbasierte Befragung standen zwei Gesichtspunkte in der folgenden Hierarchie im Fokus:

- (1) Zur Vermeidung von Designeffekten zwischen den Erhebungsmodi sollen die SAQ-Web-Fragebögen möglichst identisch mit den SAQ-Papier-Fragebögen sein, wobei die schriftlichen Fragebögen die Grundlage darstellen. Dies erlaubt eine unmittelbare Vergleichbarkeit der beiden Fragebogentypen und damit valide Trend- und Längsschnittanalysen der inhaltlichen Indikatoren.
- (2) Die Vorteile der Online-Befragung, aufgrund automatisierter Filterführung und Pflichtangaben eine gute Handhabbarkeit und möglichst geringe Belastung für die Befragten angesichts hochkomplexer Fragebögen sowie eine hohe Datenqualität zu ermöglichen, sollen optimal genutzt werden.

Die Entwicklung der schriftlichen Fragebögen stand an erster Stelle. Die Dramaturgie der Fragebogeninhalte folgte der inhaltlichen Struktur der KiGGS-Basiserhebung. Bereits verwendete Instrumente wurden identisch oder modifiziert und neue Instrumente analog der Basiserhebung umgesetzt. Nach dem Prinzip des Unified-Mode-Designs [36] wurden die Online-Fragebögen zur Vermeidung von Designeffekten beispielsweise in Bezug auf Frageformulierungen, Hervorhebungen oder Hinweistexte identisch umgesetzt. Zur optischen Unterscheidung der Single-Choice-Fragen von Multiple-Choice-Fragen wurden einheitlich für alle Erhebungsmodi die Ankreuzfelder von Single-Choice-Fragen mit runden Symbolen, die

jenigen von Multiple-Choice-Fragen mit eckigen Symbolen dargestellt. Des Weiteren wurden Multiple-Choice-Fragen zusätzlich mit dem Hinweis: »Sie können mehrere Angaben machen« gekennzeichnet. Es wurde ausschließlich die gleiche Filterführung verwendet, allerdings sind bedingt durch die technische Vorgehensweise die Filter der Online-Befragung sogenannte „harte Filter“, indem die folgenden Antworten ausgeblendet werden, was in einem schriftlichen Fragebogen nicht möglich ist. Die theoretisch mögliche Variante, lange Fragebatterien durch Filterfragen auszublenden und den Fragebogen dadurch zu verkürzen, kam nicht zum Einsatz.

Unterschiede in der Operationalisierung und Gestaltung des Online-Fragebogens nach Prinzipien des Mode-Specific-Designs [36] wurden nur dann angewandt, wenn diese nicht im Widerspruch mit der grundsätzlichen Orientierung am schriftlichen Fragebogen standen. So wurde zum Beispiel eine Drop-Down-Liste für die Abfrage des Geburtsjahres in der Web-Befragung verwendet, während diese in der Papierversion des Fragebogens mittels offener Antwortkategorie abgefragt wurde. Zur Verringerung inkonsistenter und nicht plausibler Angaben wurden, wenn möglich, zum Beispiel bei Größen- und Gewichtsangaben Plausibilitätschecks im Hintergrund programmiert, die die Befragten bei abweichenden Angaben auf nicht plausible Angaben aufmerksam machten. Des Weiteren wurden die Fragen- und Antwortitems im Rahmen der Web-Befragung so gestaltet, dass kein Scrollen für die Befragten notwendig war, sondern Frage- und Antwortitems auf einen Blick erfasst werden konnten (sogenanntes „Paging“).

Für den Fall, dass Befragte im Rahmen der Web-Befragung keine Angabe zu einer Fragestellung machten, wurden diese nachfolgend gefragt, ob sie die Angabe verweigern oder die Frage doch beantworten möchten (Do Answer Check). Durch diese Programmierung nach jeder Frage konnte ebenfalls „Satisficing“ erschwert werden.

4.6 Erfassung, Übermittlung und Sicherung der Daten

Die Datenerfassung der SAQ-Web-Befragung erfolgte mittels der Software Voxco®. Die SAQ-Papier-Fragebögen wurden mit der Software Cardiff® Teleform® erstellt, digitalisiert und verifiziert und anschließend in einem virtuellen Archiv (ELO®) gespeichert. Die beiden Programme, Voxco® und Cardiff® Teleform®, erzeugen unterschiedliche Datenstrukturen, die zur Überführung in einen Datensatz aufwändig harmonisiert werden müssen.

Die neben den Befragungsdaten vorliegenden Personen-, Verwaltungs- und Erhebungsdaten wurden in der methodischen Pilotstudie getrennt voneinander gespeichert. Zu diesem Zweck wurden separate Datenbanken und ein Pseudonymisierungsverfahren eingesetzt. Das bedeutet, dass sämtliche Daten der Verwaltungs- und Erhebungsdatenbank zwar mit Identifikationsnummern versehen sind, nicht jedoch mit Personendaten. Zudem erhalten Mitarbeiterinnen und Mitarbeiter des RKI entweder entsprechende Zugriffsrechte für die Personen- oder für die Erhebungsdaten. Durch das strikte Trennen der Daten kann ein

Tabelle 1
Eingesetzte Fragebögen in der methodischen Pilotstudie

Fragebogen	Altersgruppen		
	0–2 Jahre	3–10 Jahre	11–17 Jahre
Gesundheitsfragebogen Fassung für Eltern der 0- bis 2-Jährigen	X		
Gesundheitsfragebogen Fassung für Eltern der 3- bis 10-Jährigen		X	
Gesundheitsfragebogen Fassung für Eltern der 11- bis 17-Jährigen			X
Gesundheitsfragebogen Fassung für Jugendliche 11 bis 17 Jahre			X
Krankheitsfragebogen Fassung für Eltern der 3- bis 17-Jährigen (kurz oder lang)		X	X
Ernährungsfragebogen Fassung für Eltern der 3- bis 10-Jährigen		X	
Ernährungsfragebogen Fassung für Jugendliche 11 bis 17 Jahre			X

Zusammenfügen von Personen- und Erhebungsdaten zu keiner Zeit stattfinden.

Bei der Durchführung der Studien wurden die Bestimmungen des Bundesdatenschutzgesetzes beachtet und umgesetzt. Das Studienkonzept wurde mit Beratung des Datenschutzbeauftragten des RKI erstellt und vor Studienbeginn dem Bundesbeauftragten für den Datenschutz und die Informationsfreiheit (BfDI) vorgelegt. Die Studie wurde durch diesen geprüft und für unbedenklich erklärt. Die gegebenen Hinweise wurden berücksichtigt, gegen die Durchführung der Studie gab es keine Einwände. Die jeweiligen Landesbeauftragten für den Datenschutz der Länder, in denen die Pilotstudie durchgeführt wurde, wurden ebenfalls über die Studie informiert.

5 Ergebnisse zu Response und Repräsentativität

5.1 Vorgehensweise der methodischen Auswertungen

5.1.1 Festlegung der Falldefinition

Im Vorfeld wurden auf Basis der altersgruppenabhängigen möglichen Kombinationen von vor-

liegenden Eltern- beziehungsweise Kinder- und Jugendlichen-Fragebögen (Tabelle 2) die Bedingungen für das Erfüllen der Falldefinition festgelegt, die sowohl mit vollständigen als auch unvollständigen Befragungsunterlagen erreicht werden konnte. Für das Vorliegen eines Online-Fragebogens war definiert worden, dass dieser abgeschlossen sein musste, d. h. sich nicht im unter- oder abgebrochenen Status befand.

Ein gültiger vollständiger Fall erforderte das Vorliegen aller jeweiligen Fragebögen. In der Alterskategorie der 0- bis 2-Jährigen ist die Minimalanforderung gleichbedeutend mit der vollständigen Falldefinition, da aus inhaltlichen Gründen keine Ernährungsfragebögen ausgefüllt werden sollten. In der Altersgruppe der 3- bis 10-Jährigen mussten für einen vollständigen Fall sowohl der Gesundheitsfragebogen als auch der Ernährungsfragebogen der Eltern vorliegen. Für Teilnehmerinnen und Teilnehmer zwischen 11 und 17 Jahren erfüllte das Vorliegen des Gesundheitsfragebogens der Eltern sowie des Gesundheits- und des Ernährungsfragebogens der Kinder und Jugendlichen selbst die maximale Falldefinition. Ein gültiger unvollständiger Fall lag vor, wenn mindestens der Gesundheitsfragebogen der Eltern oder der Kinder beziehungsweise Jugendlichen vorlagen bei gleichzeitigem Fehlen mindestens eines anderen

Tabelle 2
Falldefinition für die Studienteilnahme

Altersgruppe	Gesundheitsfragebogen Eltern	Gesundheitsfragebogen Kinder und Jugendliche	Ernährungsfragebogen*	Falldefinition
0–2 Jahre	X	nicht zutreffend	nicht zutreffend	vollständig
3–10 Jahre	X	nicht zutreffend	X	vollständig
	X	nicht zutreffend	fehlt	unvollständig
	fehlt	nicht zutreffend	X	kein Fall
11–17 Jahre	X	X	X	vollständig
	X	X	fehlt	unvollständig
	X	fehlt	X	unvollständig
	fehlt	X	X	unvollständig
	X	fehlt	fehlt	unvollständig
	fehlt	X	fehlt	unvollständig
	fehlt	fehlt	X	kein Fall

* In der Altersgruppe 3 bis 10 Jahre werden die Eltern um das Ausfüllen des Ernährungsfragebogens gebeten, in der Altersgruppe 11 bis 17 Jahre die Kinder und Jugendlichen selbst. In der Altersgruppe 0 bis 2 Jahre wurde aus inhaltlichen Gründen kein Ernährungsfragebogen eingesetzt.

Fragebogens. Ein allein vorliegender Ernährungsfragebogen genügte nicht den Kriterien der Falldefinition.

5.1.2 Fallzahlen

Tabelle 3 gibt einen ersten Überblick über die Fallzahlen der Gesamtstudie, die für die detaillierten Responseberechnungen verwendet wurden, nach Vollständigkeit und Altersgruppen für alle Erhebungsdesigns. Ein Großteil der Fälle ist vollständig ($n = 3.526$, 87,4 %). In der Alterskategorie der 0- bis 2-Jährigen liegen gemäß Falldefinition keine unvollständigen Fälle vor. In den weiteren Altersgruppen, in denen lediglich die Eltern gebeten wurden, Fragebögen auszufüllen (3–6 Jahre, 7–10 Jahre), beträgt der Anteil der vollständigen Fälle 95,4 % bzw. 95,5 %. In den Altersgruppen, in denen auch Kinder und Jugendliche Fragebögen beantworten sollten, ist die Anzahl der vollständigen Fälle niedriger (11–13 Jahre: 71,6 %, 14–17 Jahre: 75,4 %). In 66,2 % aller unvollständigen Fälle haben dabei ausschließlich die Eltern den Gesundheitsfragebogen ausgefüllt und in 18,0 % nur die Kinder und Jugendlichen.

5.2 Teilnahme (Response)

5.2.1 Teilnahme nach Erhebungsdesign

Für die Berechnung der Responsequoten wurde für jedes Erhebungsdesign die Anzahl der Teilnehmerinnen und Teilnehmer durch den Umfang der Bruttostichprobe abzüglich der qualitätsneutralen Ausfälle (bereinigte Bruttostichprobe) dividiert. Dieses Vorgehen entspricht dem international anerkannten Standard der „American Association for Public Opinion Research“ (AAPOR) von „Response Rate 2“ [149].

$$\text{Response Rate} = \frac{\text{Teilnehmer}}{\text{Bruttostichprobe} - \text{QNA}}$$

Die Bruttostichprobe der Gesamtstudie bestand aus insgesamt 11.140 Kindern und Jugendlichen im Alter von 0 bis 17 Jahren. Aus dieser Bruttostichprobe wurden zunächst die qualitätsneutralen Ausfälle (QNA) ausgeschlossen, also Personen, die als nicht der Zielpopulation zugehörig definiert wurden. Als Gründe hierfür war festgelegt worden, dass die Personen aufgrund einer unbekannt oder falschen Adresse nicht auffindbar sind, eine sprachliche Verständigung nicht möglich ist, das Alter außerhalb des zuvor festgelegten Altersbereichs liegt, die Personen verzogen sind, hauptsächlich in einer Wohnung außerhalb des Sample Points leben, doppelt gezogen wurden oder verstorben sind. In der Pilotstudie KiGGS Welle 2 traten insgesamt 647 QNA auf, die sich anteilig gleichmäßig auf die Erhebungsdesigns verteilen (Tabelle 4). Insgesamt beteiligten sich 4.032 Personen mit gültigen Befragungsunterlagen an der Studie.

Auf dieser Grundlage konnten die Responsequoten für die verschiedenen Erhebungsdesigns ermittelt werden. In der Kontrollgruppe, in der ausschließlich die schriftliche Befragung angeboten wurde, fand sich eine Responsequote von 42,0 %. In der simultanen Mixed-Mode-Gruppe mit gleichzeitigem Angebot des schriftlichen und webbasierten Fragebogens lag die Responsequote bei 41,4 %. Im sequenziellen Mixed-Mode-Design wurde eine Responsequote von 38,5 % und im Preselect-Design von 30,7 % erzielt (Tabelle 4).

Der Vergleich der Responsequoten der jeweiligen Erhebungsdesigns zeigt, dass durch ein zusätzliches SAQ-Web-Teilnahmeangebot unabhängig von der Abfolge, in der dieses Angebot im Verhältnis zum SAQ-Papier-Teilnahmeangebot unterbreitet wurde, keine zusätzlichen Teilnehmerinnen und Teilnehmer gewonnen werden konnten.

Tabelle 3
Vollständige und unvollständige Fälle nach Altersgruppen

	0–2 Jahre	3–6 Jahre	7–10 Jahre	11–13 Jahre	14–17 Jahre	Gesamt
Vollständiger Fall	571	908	863	502	682	3.526
Unvollständiger Fall	0	44	41	199	222	506
Gesamt	571	952	904	701	904	4.032

Die Differenz zwischen dem Preselect- und den anderen Erhebungsdesigns ist ähnlich hoch, wenn statt der tatsächlichen Teilnahme die Teilnahmebereitschaft betrachtet wird. Als Teilnahmebereitschaft ist definiert, wenn Personen bekundeten, dass sie teilnehmen würden, zum Beispiel im Rahmen der telefonischen Werbung, dies aber gegebenenfalls letztlich doch nicht taten. Die Teilnahmebereitschaft beträgt in der Kontrollgruppe 45,3 %, im sequenziellen Erhebungsdesign 42,0 %, im simultanen Erhebungsdesign 44,4 % und im Preselect-Erhebungsdesign 34,5 %. Es zeigt sich, dass die im letztgenannten Design getestete Verfahrensweise, wonach die jeweils gewünschten Erhebungsunterlagen erst auf individuellen Wunsch bereitgestellt werden, zu keinem übermäßigen Verlust führt, sondern von vornherein hier eine niedrigere Quote hinsichtlich der Teilnahmebereitschaft erzielt wird.

Die Betrachtung nach Altersgruppen und Erhebungsdesigns zeigt, dass im sequenziellen und

simultanen Mixed-Mode-Design die höchsten Responsequoten tendenziell in der jüngsten Altersgruppe und die niedrigsten Responsequoten in der höchsten Altersgruppe erzielt wurden (Tabelle 5). Dagegen wurden in der Kontrollgruppe im Mixed-Mode preselect ähnlich hohe Responsequoten über alle Altersgruppen erreicht.

Die Ergebnisse bestätigen die vorliegenden Forschungsbefunde, die es fraglich erscheinen lassen, ob durch Mixed-Mode-Designs mit SAQ-Papier- und SAQ-Web-Teilnahmeangebot höhere Responsequoten zu erzielen sind als bei einer ausschließlichen Teilnahmemöglichkeit mit schriftlichen Fragebögen (vgl. Kapitel 3.1). Der bevölkerungsbezogene Ansatz sowie die sensiblen Themen der KiGGS-Studie bilden dabei allerdings einen Rahmen, der Vergleiche zu anderen Studien generell erschwert. Beispielweise konnte im deutschen Epidemiologischen Suchtsurvey durch den Einsatz eines zusätzlichen SAQ-Web-Teilnahmeangebots die Responsequote gesteigert werden

Tabelle 4
Responseberechnung für die methodische Pilotstudie

	Single- Mode	Mixed- Mode sequenziell	Mixed- Mode simultan kurz	Mixed- Mode simultan lang	Mixed- Mode preselect	Gesamt
Bruttostichprobe	2.477	2.475	2.474	1.243	2471	11.140
Teilnehmende	979	899	965	472	717	4.032
Unvollständige Fälle	111	146	129	56	64	506
Vollständige Fälle	868	753	836	416	653	3.526
Nichtteilnehmende	1.353	1.434	1.368	687	1.619	6.461
Kein Kontakt	951	976	943	449	1.055	4.374
Nur Nonresponder-Fragebogen beantwortet	211	234	234	124	267	1.070
Verweigerung (abgesagt)	85	93	81	57	112	428
Nur Teilnahmebereitschaft bekundet	78	81	78	42	88	367
Einwilligung unterschrieben, aber keine Fragebögen	19	22	23	7	67	138
Fragebögen vorhanden, aber keine Einwilligung unterschrieben	0	0	0	0	0	0
Teilnahme nicht gemäß Falldefinition (nur Ernährungsfragebogen)	1	7	3	2	7	20
Sonstige	8	21	6	6	23	64
Qualitätsneutrale Ausfälle (QNA)	145	142	141	84	135	647
Bereinigte Bruttostichprobe	2.332	2.333	2.333	1.159	2.336	10.493
Responsequote	42,0 %	38,5 %	41,4 %	40,7 %	30,7 %	38,5 %

Tabelle 5
Responsequoten nach Altersgruppen und Erhebungsdesigns

	Single-Mode	Mixed-Mode sequenziell	Mixed-Mode simultan kurz	Mixed-Mode preselect	Gesamt
	%	%	%	%	
0–2 Jahre	46,6	43,7	48,0	29,0	41,7
3–6 Jahre	42,5	42,5	42,1	31,0	39,6
7–10 Jahre	39,4	38,7	43,9	29,2	37,8
11–13 Jahre	43,6	39,7	43,1	36,5	40,8
14–17 Jahre	40,4	31,4	33,9	28,8	33,6

[58]. Allerdings wurde der Methodenmix anders umgesetzt und bezog sich auf die Erwachsenenbevölkerung im Alter von 18 bis 64 Jahren [58]. Die niedrigere Responsequote im simultanen Mixed-Mode-Design der vorliegenden Pilotstudie von KiGGS Welle 2 bestätigt die Ergebnisse der Meta-Analyse von Medway und Fulton (2012) [40]. Die Hypothese, dass im Rahmen der Pilotstudie möglicherweise aufgrund der relativ jungen Zielpopulation von Eltern, Kindern und Jugendlichen höhere Responsequoten durch den Einsatz des SAQ-Web-Teilnahmeangebotes erzielt werden könnten, kann nicht bestätigt werden. Hingegen zeigen sich ähnliche Ergebnisse wie bei der Nord-Child-Studie in Dänemark aus dem Jahr 2011, bei der das Single-Mode-Design mit schriftlichen Fragebögen die höchste Responsequote lieferte [65].

5.2.2 Teilnahme nach Erhebungsmodus innerhalb der Erhebungsdesigns

Mit der Überprüfung der Nutzungshäufigkeit der verschiedenen Erhebungsmodi, die in den Mixed-Mode-Erhebungsdesigns angeboten wurden, kann die Präferenz für einen Erhebungsmodus gegenüber dessen Toleranz überprüft werden. Die Nutzung eines Erhebungsmodus bei der freien Auswahl (wie in den simultanen Mixed-Mode-Designs sowie im Preselect-Design) kann als Präferenz, die Nutzung des Modus bei dessen ausschließlichem Angebot (wie zunächst beim Webangebot im sequenziellen Design) jedoch nur als Toleranz interpretiert werden [24].

Werden bestimmte Erhebungsmodi generell oder unter bestimmten Angebotsabfolgen nicht genutzt, ist das entsprechende Verfahren nicht

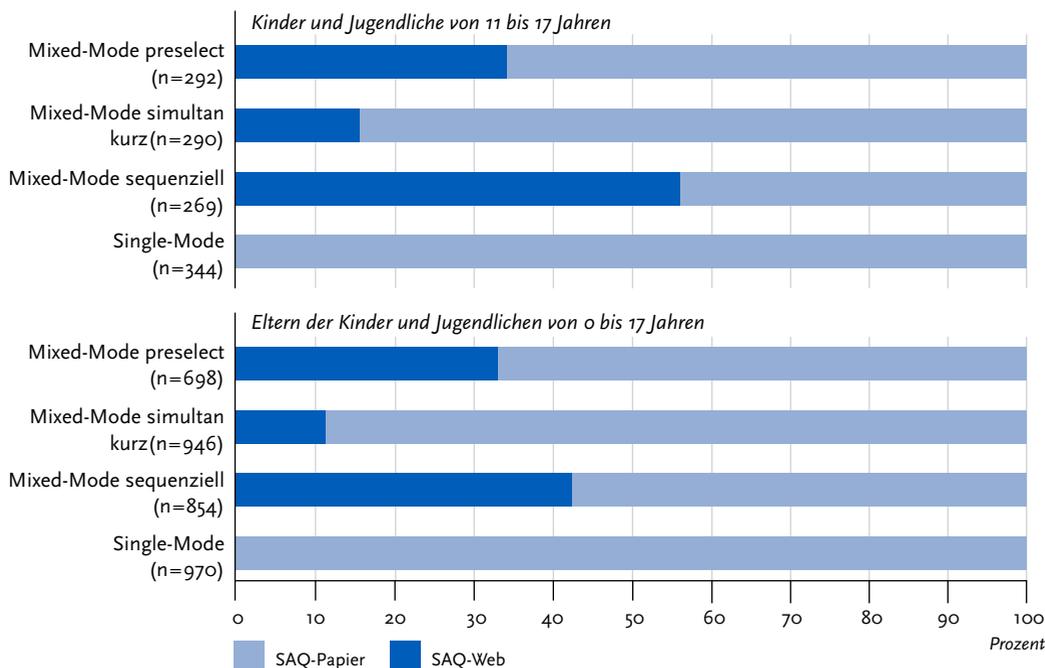
zweckdienlich [24]. Mit Blick auf die methodische Pilotstudie ist die Frage relevant, ob der zusätzlich angebotene Online-Modus von der Studienpopulation der KiGGS-Studie bevorzugt gewählt wird.

Insgesamt haben im Web-Modus, dem zusätzlich zur schriftlichen Erhebung angebotenen Erhebungsmodus, 760 Eltern und 323 Kinder und Jugendliche den Gesundheitsfragebogen ausgefüllt. Am häufigsten wird die Web-Befragung durch fast die Hälfte aller Eltern im sequenziellen Erhebungsdesign genutzt (42,9 %), im Preselect-Erhebungsdesign durch etwa ein Drittel (33,4 %) und im simultanen Erhebungsdesign von lediglich 11,5 % der teilnehmenden Eltern (Abbildung 5). Die Anteile für die online teilnehmenden Kinder und Jugendlichen in den verschiedenen Erhebungsdesigns unterscheiden sich insbesondere in der simultanen und Preselect-Gruppe kaum von jenen der Eltern. Einzig im sequenziellen Erhebungsdesign verkehrt sich das Verhältnis der Web- und Papier-Modus-Nutzung, indem mit 56,5 % eine knappe Mehrheit der Kinder und Jugendlichen im Web-Modus teilnimmt.

Damit wird deutlich, dass das Web-Angebot in allen Mixed-Mode-Erhebungsdesigns zwar genutzt wird, dass aber hinsichtlich der relativen Häufigkeit die Abfolge der Angebote relevant ist. Der höhere Anteil der Web-Nutzung im sequenziellen Erhebungsdesign ist darauf zurückzuführen, dass hier – anders als in den Erhebungsdesigns mit parallelem Angebot verschiedener Befragungsformen – zunächst ausschließlich die Möglichkeit der Web-Teilnahme unterbreitet wird. Im Vergleich des sequenziellen mit den parallelen Erhebungsdesigns wird deutlich, dass damit lediglich eine Toleranz gegenüber der Web-Teilnahme besteht, diese Erhebungsform im Vergleich zu einer Teil-

Abbildung 5

Anteil der Erhebungsmodi bei Kindern und Jugendlichen sowie bei den Eltern nach Erhebungsdesign*



* Es werden die für den Gesundheitsfragebogen genutzten Modi betrachtet. Abweichungen zu den Teilnehmerzahlen in Tabelle 4 sind entsprechend der Falldefinition möglich (siehe Tabelle 2).
SAQ=Self-Administered Questionnaire

nahme per schriftlichen Fragebögen jedoch nicht präferiert wird. [24]. Diese Tatsache bestätigt bereits vorliegende Ergebnisse, wonach ein relativ hoher Anteil von Web-Teilnehmerinnen und -Teilnehmern nur in einem sequenziellen Mixed-Mode-Design mit dem Online-Modus als erstem Angebot erzielt werden kann [37, 41]. Dass die Kinder und Jugendlichen anders als erwartet die Web-Befragung nicht durchgehend häufiger wählen als die Eltern, ist allerdings nicht unbedingt eine Frage der Präferenz, sondern könnte auch darin begründet sein, dass eine innerfamiliäre, einheitliche Entscheidung für eine Befragungsform mit oder durch die Eltern getroffen wird.

Grundsätzlich bestand für Eltern sowie Kinder und Jugendliche innerhalb einer Familie die Möglichkeit, unterschiedliche Befragungsformen zu nutzen. In der Altersgruppe der 3- bis 10-Jährigen konnten auch die Eltern selbst im sequenziellen (nach der Erinnerung) und simultanen Erhebungsdesign zwischen den Erhebungsmodi wech-

seln, was allerdings nie geschah. Bei Kindern und Jugendlichen konnte dies nicht vorkommen, weil ihnen die Fragebögen stets nur in einem Modus bereitgestellt wurden. In 8,1% der Fälle füllten Eltern und Kinder beziehungsweise Jugendliche ihren Gesundheitsfragebogen in unterschiedlichen Erhebungsmodi aus. Dabei nutzten in 25,1% der Fälle die Kinder beziehungsweise Jugendlichen das Web-Angebot, während die Eltern die schriftliche Variante nutzten. Andersherum waren es nur 1,1%. Nach Erhebungsdesign betrachtet findet sich der höchste Anteil an „gemischten Erhebungen“ im sequenziellen Erhebungsdesign mit 10,7%, gefolgt vom simultanen Erhebungsdesign mit 8,9% und dem Preselect-Erhebungsdesign mit 5,5%. Diese insgesamt niedrigen Anteile zeigen, dass eine theoretisch diskutierte höhere Flexibilität und damit vermeintlich höhere Nutzerfreundlichkeit von Mixed-Mode-Designs im Sinne individueller Präferenzen von Erhebungsformen hier nur wenig praxisrelevant ist.

5.2.3 Nichtteilnahme (Nonresponse)

Eine Nichtteilnahme ist gemäß der Falldefinition (vgl. Kapitel 5.1.1) gegeben, wenn kein Gesundheitsfragebogen (weder der Eltern noch der Kinder beziehungsweise Jugendlichen) vorliegt. Qualitätsneutrale Ausfälle werden dabei nicht berücksichtigt, weil sie aus der bereinigten Bruttostichprobe ausgeschlossen sind. Insgesamt beteiligten sich 6.461 Personen nicht an der Studie. Zum weitaus größten Teil an Nicht-Teilnehmenden ($n=4.374$, 67,3 %), sowohl insgesamt als auch zu ähnlichen Anteilen innerhalb aller Erhebungsdesigns, konnte keine endgültige Information zur Teilnahmebereitschaft (d. h. weder eine Zu- noch Absage) ermittelt werden (vgl. Tabelle 4). Dies ist vermutlich auf die im Vergleich zur KiGGS-Hauptstudie reduzierten Werbungsmaßnahmen zurückzuführen, wobei insbesondere eine eingeschränkte Presse- und Öffentlichkeitsarbeit sowie die Unterlassung von Hausbesuchen zur Gewinnung von TeilnehmerInnen und Teilnehmern eine Rolle spielen. In 1.070 Fällen wurde ein Nonresponder-Fragebogen ausgefüllt. Damit liegen für insgesamt 10,2 % der bereinigten Bruttostichprobe Informationen wenigstens zu einigen zentralen Fragestellungen vor. Dagegen lehnten nur 4,0 % ($n=428$) auch die Beantwortung des Nonresponder-Fragebogens ab. Weitere Nichtteilnehmerinnen und Nichtteilnehmer hatten ihre Teilnahmebereitschaft bekundet ($n=367$, 3,4 %), realisierten diese aber letztlich nicht. Darüber hinaus lag von 134 Personen (1,3 %) eine unterschriebene Einwilligungserklärung vor, ohne dass ausgefüllte Fragebögen vorhanden gewesen wären. Diese niedrigen Zahlen deuten zum einen darauf hin, dass verwaltungstechnischer Aufwand seitens der eingeladenen Personen keine Hürde für die Teilnahme darstellte und auch die Online-Erhebungsform nicht zu einer relevanten Menge an „Abbrüchen“ führte. Zum anderen legen sie nahe, dass durch systematische Follow-ups die Anzahl an Teilnehmerinnen und Teilnehmern oder wenigstens die Anzahl derer, die einen Nonresponder-Fragebogen beantworteten, erhöht werden kann.

5.3 Stichprobenzusammensetzung

Die Responsequote sowie die Häufigkeit der Nutzung der verschiedenen Erhebungsmodi sind

nur zwei Kriterien, mit denen der Nutzen eines Mixed-Mode-Designs im Vergleich zu einem Single-Mode-Design beurteilt werden kann. Eine weitere Fragestellung der Studie ist, ob eine mit einem Mixed-Mode-Design realisierte Stichprobe besser dazu in der Lage ist, die Zielpopulation in ihren soziodemografischen Merkmalen zu repräsentieren. Dafür müssten andere Personengruppen mit den verschiedenen Teilnahmeformen erreicht werden und sich damit die Zusammensetzung der Stichproben unterscheiden.

5.3.1 Stichprobenzusammensetzung der Erhebungsdesigns

Für die aus den Einwohnermeldeämtern bekannten soziodemografischen Merkmale Geschlecht, Alter, Staatsangehörigkeit, Gemeindegröße und Region wurden zunächst die Bruttostichproben der verschiedenen Erhebungsdesigns miteinander verglichen (siehe Tabelle 6, kursiv). Es sind keine Unterschiede zwischen den Designs vorhanden (p -Werte nicht dargestellt), wodurch gezeigt werden kann, dass die Verteilung der Gesamtstichprobe auf die einzelnen Erhebungsdesigns zufällig ist. Vor diesem Hintergrund können Unterschiede in der Zusammensetzung der einzelnen realisierten Stichproben als „nicht zufällig“ eingestuft werden.

Anschließend wurden die realisierten Stichproben eines jeden Mixed-Mode-Designs mit der Single-Mode-Kontrollgruppe in ihrer soziodemografischen Zusammensetzung auf Unterschiedlichkeit geprüft (Tabelle 6). Dafür wurden ebenfalls die aus den Registern bekannten Informationen sowie aus Erhebungsdaten gewonnene Informationen, wie zum Beispiel Migrationshintergrund, Bildung oder Einkommen, herangezogen. Die statistische Prüfung erfolgte mittels χ^2 -Test. Das Signifikanzniveau wurde für multiples Testen mit der Bonferroni-Korrektur auf $p < 0,017$ adjustiert [32], da insgesamt jeweils drei statistische Tests bei jeder soziodemografischen Eigenschaft durchgeführt wurden. Die Wahrscheinlichkeit eines zufällig signifikanten Ergebnisses sollte dadurch verringert werden.

Wie Tabelle 6 zeigt, sind nach Korrektur des Signifikanzniveaus für multiples Testen keine statistisch signifikanten Unterschiede in der soziodemografischen Zusammensetzung zwischen einem der Mixed-Mode-Designs und der Single-Mode-

Tabelle 6

Stichprobenzusammensetzung der realisierten Stichproben und der Bruttostichproben nach Erhebungsdesign

	Single-Mode-Design		Sequenzielles Mixed-Mode-Design			Simultanes Mixed-Mode-Design kurz			Preselect Mixed-Mode-Design		
	Teilnehmende	Bruttostichprobe*	Teilnehmende	Bruttostichprobe*	p-Wert**	Teilnehmende	Bruttostichprobe*	p-Wert**	Teilnehmende	Bruttostichprobe*	p-Wert**
	%	%	%	%		%	%		%	%	
Geschlecht des Kindes											
Weiblich	52,3	49,8	47,6	47,4		52,6	50,3		49,1	49,2	
Männlich	47,7	50,2	52,4	52,6	0,043	47,4	49,8	0,879	50,9	50,8	0,192
Alter des Kindes											
0–2 Jahre	14,2	12,8	14,4	12,6	0,313	15,0	12,9	0,287	12,6	13,3	0,734
3–6 Jahre	22,9	22,6	25,4	23,0		23,1	22,7		22,5	22,3	
7–10 Jahre	21,1	22,6	22,9	22,8		24,0	22,7		21,5	22,6	
11–13 Jahre	17,5	16,8	16,7	16,2		17,2	16,5		19,7	16,5	
14–17 Jahre	24,3	24,3	20,7	25,4		20,6	25,2		23,9	25,4	
Erste Staatsangehörigkeit											
Deutsch	90,4	86,8	90,2	86,8	0,425	88,8	86,4	0,154	90,0	87,4	0,310
Nicht-Deutsch	9,6	13,2	9,8	13,2		11,2	13,6		10	12,6	
Urbanisierungsgrad											
Ländlich/kleinstädtisch (<20.000)	47,4	47,5	49,8	47,7	0,485	49,3	48,1	0,696	49,7	49,7	0,270
Mittelstädtisch (20.000–<100.000)	22,8	24,0	20,8	23,0		22,0	23,2		19,5	21,8	
Großstädtisch (≥100.000)	29,8	28,5	29,4	29,2		28,7	28,6		30,8	28,5	
Region											
Westdeutschland	52,8	50,3	50,3	49,6	0,273	51,4	50,7	0,534	53,3	50,8	0,849
Ostdeutschland/Berlin	47,2	49,7	49,7	50,4		48,6	49,3		46,7	49,2	
Migrationshintergrund											
Ein- oder beidseitig	11,4	N/A	11,9	N/A	0,920	12,1	N/A	0,895	10,1	N/A	0,607
Interviewter Elternteil											
Mutter	78,5	N/A	78,4	N/A	0,196	80,4	N/A	0,396	79,3	N/A	0,075
Vater	6,2	N/A	8,4	N/A		6,9	N/A		6,9	N/A	
Mutter und Vater gemeinsam	14,4	N/A	12,4	N/A		12,1	N/A		13,8	N/A	
Andere	0,93	N/A	0,70	N/A		0,64	N/A		0,00	N/A	

Fortsetzung nächste Seite

* Vergleich der Bruttostichproben der verschiedenen Erhebungsdesigns miteinander

** Entsprechendes Mixed-Mode-Design verglichen mit dem Single-Mode-Design

N/A=Informationen über die Bruttostichprobe sind nicht vorhanden

Tabelle 6 Fortsetzung

Stichprobenzusammensetzung der realisierten Stichproben und der Bruttostichproben nach Erhebungsdesign

	Single-Mode-Design		Sequenzielles Mixed-Mode-Design			Simultanes Mixed-Mode-Design kurz			Preselect Mixed-Mode-Design		
	Teilnehmende	Bruttostichprobe*	Teilnehmende	Bruttostichprobe*	p-Wert**	Teilnehmende	Bruttostichprobe*	p-Wert**	Teilnehmende	Bruttostichprobe*	p-Wert**
	%	%	%	%		%	%		%	%	
Familienstand											
Verheiratet, zusammenlebend	71,9	N/A	70,1	N/A	0,031	70,5	N/A	0,520	71,6	N/A	0,148
Getrennt/geschieden/verwitwet	12,0	N/A	9,6	N/A		11,4	N/A		9,6	N/A	
Ledig	16,1	N/A	20,3	N/A		18,1	N/A		18,8	N/A	
Bildung des Haushaltsalts^a											
Untere Bildungsgruppe	7,8	N/A	7,8	N/A	0,941	8,5	N/A	0,640	6,5	N/A	0,043
Mittlere Bildungsgruppe	56,8	N/A	57,5	N/A		57,9	N/A		52,2	N/A	
Obere Bildungsgruppe	35,4	N/A	34,7	N/A		33,6	N/A		41,4	N/A	
Haushaltsnettoeinkommen^b											
Niedrig (Q1)	20,2	N/A	23,0	N/A	0,120	21,4	N/A	0,831	18,6	N/A	0,346
Mittel (Q2–4)	65,4	N/A	60,5	N/A		64,5	N/A		64,4	N/A	
Hoch (Q5)	14,4	N/A	16,5	N/A		14,1	N/A		17,0	N/A	

* Vergleich der Bruttostichproben der verschiedenen Erhebungsdesigns miteinander

** Entsprechendes Mixed-Mode-Design verglichen mit dem Single-Mode-Design

^a CASMIN-Bildungsgruppen; CASMIN = Comparative Analysis of Social Mobility in Industrial Nations^b Äquivalenzeinkommen nach neuer OECD-Skala, Median nach EU-SILC 2010; OECD = Organisation for Economic Co-operation and Development;

EU-SILC = European Union Statistics on Income and Living Conditions

N/A = Informationen über die Bruttostichprobe sind nicht vorhanden

Kontrollgruppe vorhanden. Die Ergebnisse zeigen, dass es mit keinem der angebotenen Mixed-Mode-Designs gelingt, Teilnehmerinnen und Teilnehmer mit anderen Eigenschaften zu gewinnen und eine bessere Repräsentativität der Stichprobe zu erreichen. Dieses Ergebnis steht im Einklang mit bisherigen Erkenntnissen (vgl. Kapitel 3.2).

5.3.2 Stichprobenzusammensetzung der Modegruppen

Der Vergleich der soziodemografischen Merkmale nach Erhebungsmodus erfolgte sowohl für die Gruppe der Eltern als auch für die Gruppe der Kinder und Jugendlichen von 11 bis 17 Jahren. Einbezogen wurden Eigenschaften des Kindes (Geschlecht, Alter, höchster vorhandener oder

angestrebter Schulabschluss, Migrationshintergrund), des befragten Elternteils (Familienstand), des Haushaltes, in dem das Kind lebt (höchste Bildung im Haushalt nach CASMIN-Bildungsgruppen, Einkommen) und der Gemeinde (Gemeindegröße, Regionalität).

Gesundheitsfragebögen zu Kindern und Jugendlichen im Alter von 0 bis 17 Jahren liegen von insgesamt 3.936 Eltern vor. Selbstangaben liegen für insgesamt 1.340 Kinder und Jugendliche im Alter von 11 bis 17 Jahren vor. Für den Vergleich zwischen den Erhebungsmodi wurde das Single-Mode-Design nicht berücksichtigt, sondern lediglich als externe Vergleichsgruppe betrachtet, da die Teilnehmerinnen und Teilnehmer keine Wahlmöglichkeit bezüglich des Erhebungsmodus hatten. Die statistische Prüfung auf Unterschiede erfolgte mit dem Chi²-Test auf dem Signifikanzniveau von $p < 0,05$.

Tabelle 7

Soziodemographische Zusammensetzung der realisierten Stichproben nach Erhebungsmodus
(getrennt nach Kontrollgruppe und Mixed-Mode-Gruppen)

	Proxyangaben der Eltern (alle Altersgruppen)				Selbstangaben der Kinder und Jugendlichen (11–17 Jahre)			
	Kontrollgruppe		Mixed-Mode-Designs		Kontrollgruppe		Mixed-Mode-Designs	
	SAQ-Papier n=970	SAQ-Web n=760	SAQ-Papier n=2.206	p-Wert	SAQ-Papier n=343	SAQ-Web n=323	SAQ-Papier n=674	p-Wert
	%	%	%		%	%	%	
Geschlecht des Kindes								
Weiblich	52,5	47,5	50,1	0,210	53,9	45,5	53	0,028
Männlich	47,5	52,5	49,9		46,1	54,5	47	
Alter des Kindes								
0–2 Jahre	14,3	15,1	14,4		–	–	–	
3–6 Jahre	23,1	27,8	23,4		–	–	–	
7–10 Jahre	21,3	25,8	22,7		–	–	–	
11–13 Jahre	17,3	15,5	17,0		40,5	44,5	44,5	
14–17 Jahre	23,9	15,8	22,5	0,001	59,9	55,4	55,5	0,085
Migrationshintergrund								
Ein- oder beidseitig	11,2	11,6	10,9	0,113	12,2	13,0	11,3	0,631
Region								
Ost (inkl. Berlin)	52,8	49,1	53,4		52,5	48,3	53,3	
West	47,2	50,9	46,6	0,042	47,5	51,7	46,8	0,142
Wohnort								
Ländlich/kleinstädtisch (<20.000)	47,3	47,5	49,5		47,8	51,1	49,3	
Mittelstädtisch (20.000–<100.000)	22,8	19,6	21,3		21,6	21,4	20,8	
Großstädtisch (≥100.000)	29,9	32,9	29,1	0,142	30,6	27,6	30,0	0,743
Bildung des Haushaltsalts^a								
Nicht ermittelbar	0,3	0	0,3		0,6	0	0	
Untere Bildungsgruppe	7,7	4,1	8,9		8,4	1,9	11,0	
Mittlere Bildungsgruppe	56,6	50,5	58,7		58,4	56,9	59,1	
Obere Bildungsgruppe	35,4	45,4	32,0	0,000	32,6	41,2	29,9	0,000
Haushaltsnettoeinkommen^b								
Fehlende Angaben	10,7	0	12,5		12	6,3	12,2	
Niedrig (Q1)	18,0	15,7	21,2		16,2	13,9	20,1	
Mittel (Q2–4)	58,5	58,6	54,4		56,6	55,8	55,5	
Hoch (Q5)	12,9	20,8	11,9	0,000	15,3	24,0	11,8	0,000

Fortsetzung nächste Seite

^a CASMIN-Bildungsgruppen; CASMIN=Comparative Analysis of Social Mobility in Industrial Nations

^b Äquivalenzeinkommen nach neuer OECD-Skala, Median nach EU-SILC 2010; OECD=Organisation for Economic Co-operation and Development;

EU-SILC=European Union Statistics on Income and Living Conditions

SAQ=Self-Administered Questionnaire

Angaben, die sich nicht auf 100 summieren, gehen auf unplausible oder fehlende Angaben zurück

p-Wert=Signifikanztest zwischen den Erhebungsmethoden innerhalb der Mixed-Mode Designs

Tabelle 7 Fortsetzung

Soziodemographische Zusammensetzung der realisierten Stichproben nach Erhebungsmodus (getrennt nach Kontrollgruppe und Mixed-Mode-Gruppen)

	Proxyangaben der Eltern (alle Altersgruppen)				Selbstangaben der Kinder und Jugendlichen (11–17 Jahre)			
	Kontrollgruppe		Mixed-Mode-Designs		Kontrollgruppe		Mixed-Mode-Designs	
	SAQ-Papier n=970	SAQ-Web n=760	SAQ-Papier n=2.206	p-Wert	SAQ-Papier n=343	SAQ-Web n=323	SAQ-Papier n=674	p-Wert
	%	%	%		%	%	%	
Bildung des Kindes								
Unbekannt					4,4	8,4	5,9	
Niedriger Schulabschluss angestrebt oder vorh.	–	–	–		5,0	2,5	5,9	
Mittlerer Schulabschluss angestrebt oder vorh.	–	–	–		27,7	22,0	29,5	
Hoher Schulabschluss angestrebt oder vorh.	–	–	–		63,0	67,2	58,6	0,002
Interviewpartnerin/Interviewpartner								
Mutter	78,1	73,9	81,4		–	–	–	–
Vater	6,2	11,4	5,9		–	–	–	–
Mutter und Vater gemeinsam	14,3	14,3	12,2	0,000	–	–	–	–
Familienstand der Interviewpartnerin/des Interviewpartners								
Verheiratet, zusammenlebend	71,1	74,3	68,7		–	–	–	–
Getrennt/geschieden/verwitwet	11,9	8,3	11,6		–	–	–	–
Ledig	16,0	17,4	19,6	0,001	–	–	–	–
Alter der Interviewpartnerin/des Interviewpartners								
≤30 Jahre	11,3	11,0	11,4		–	–	–	–
31–40 Jahre	42,4	45,8	43,3		–	–	–	–
41–50 Jahre	40,2	39	40,7		–	–	–	–
<50 Jahre	6,0	4,2	4,7	0,692	–	–	–	–

SAQ=Self-Administered Questionnaire

Angaben, die sich nicht auf 100 summieren, gehen auf unplausible oder fehlende Angaben zurück

p-Wert=Signifikanztest zwischen den Erhebungsmethoden innerhalb der Mixed-Mode Designs

Tabelle 7 verdeutlicht, dass in den Mixed-Mode-Gruppen die online teilnehmenden Eltern häufiger männlich sind als diejenigen, die den schriftlichen Fragebogen beantwortet haben. Sie haben jüngere Kinder, ein höheres Einkommen und eine höhere Bildung, sind häufiger verheiratet und leben häufiger in den alten als in den neuen Bundesländern.

Die Stichprobenszusammensetzung bei den Kindern und Jugendlichen in den Mixed-Mode-Grup-

pen unterscheidet sich ebenfalls zwischen den beiden Erhebungsmodi. Angesichts geringerer Fallzahlen sind aber weniger signifikante Unterschiede festzustellen. Zusammenfassend kann hier für die Mixed-Mode-Erhebungsdesigns gesagt werden, dass online teilnehmende Kinder und Jugendliche ebenfalls häufiger männlich sind. Sie haben einen höheren Schulabschluss beziehungsweise streben einen höheren Schulabschluss an und ihre Eltern

besitzen eine höhere Bildung sowie ein höheres Einkommen als Teilnehmerinnen und Teilnehmer, die den schriftlichen Fragebogen genutzt haben.

Der Blick auf Unterschiede zwischen den Erhebungsmodi deutet im Einklang mit bisherigen Forschungsergebnissen (vgl. Kapitel 3.2) darauf hin, dass über das Angebot eines Online-Fragebogens vor allem höhere Anteile an Personengruppen mit einem höheren sozialen Status erreicht werden können, die sich in der Vergangenheit ohnehin als stärker teilnahmebereit gezeigt haben als andere Gruppen. Die fehlenden Unterschiede in der regionalen Verteilung zwischen den Modegruppen zeigen auf, dass keine Probleme bestehen, ländliche Regionen mit einem SAQ-Web-Teilnahmeangebot zu erreichen. Allerdings werden Teilnehmerinnen und Teilnehmer in Großstädten gleichzeitig auch nicht besser angesprochen, was wünschenswert wäre, weil sie als schwer erreichbar gelten [26].

Insgesamt können die angeführten Betrachtungen vorsichtig als konservativ eingeschätzt werden, weil die Teilnehmerinnen und Teilnehmer des sequenziellen Erhebungsdesigns, welche sich für das Online-Teilnahmeangebot entschieden ohne von dem Papier-Teilnahmeangebot zu wissen, hier der Web-Teilnehmergruppe zugeordnet wurden. Da auf Basis der Erkenntnisse zum simultanen Mixed-Mode-Design angenommen werden kann, dass sie auch das Papier-Teilnahmeangebot angenommen hätten, unterscheiden sich die Teilnehmergruppen zwischen SAQ-Papier- und SAQ-Web-Teilnahmeform tendenziell weniger als hier dargestellt.

6 Ergebnisse zu Modeunterschieden zwischen SAQ-Papier und SAQ-Web

6.1 Item-Nonresponse und Datenqualität

Der Einsatz von Mixed-Mode-Designs hat Einfluss auf die Datenqualität, insbesondere die Aussagekraft und Vergleichbarkeit können durch den Einsatz unterschiedlicher Erhebungsmodi betroffen sein. Die Analyse der Item-Nonresponse, also fehlender Werte aufgrund von Antwortverweigerung einzelner Frageitems, stellt eine Möglichkeit zur Beurteilung der Datenqualität der beiden eingesetzten Erhebungsmodi SAQ-Papier und SAQ-Web dar. Item-Nonresponse führt zu einem Informationsverlust. Dadurch können Berechnungen

ungenauer werden und statistische Testverfahren weisen eine geringere Power auf [102].

6.1.1 Methodik

In der Pilotstudie von KiGGS Welle 2 wurden im Rahmen der Datenaufbereitung fehlende Werte definiert und verschiedenen Kategorien zugeordnet. In die Analysen zur Item-Nonresponse werden nur diejenigen fehlenden Werte einbezogen, die aufgrund von Antwortverweigerung oder dem Auslassen einer einzelnen Antwortkategorie bei Mehrfachangaben bei eigentlich von dem jeweiligen Teilnehmenden zu beantwortenden Fragen entstehen („keine Angabe“). Sogenannte filterbedingte Missings, also fehlende Angaben vor dem Hintergrund des Nicht-Zutreffens der Frage aus verschiedenen Gründen, werden nicht mit betrachtet. In der KiGGS-Studie wird beispielsweise eine Vielzahl von Fragen nur bestimmten Altersgruppen oder nur den Mädchen beziehungsweise Jungen gestellt oder vertiefende Fragen müssen nur beantwortet werden, wenn eine Eingangsfrage mit „ja“ beantwortet wird. Die Antwortmöglichkeit „weiß nicht“, wird in diesem Bericht auch als fehlender Wert behandelt, da Teilnehmende keine passende Antwortkategorie gefunden kann.

Zudem wurden in der Pilotstudie nur vollständig ausgefüllte Online-Fragebögen in die Analysen mit einbezogen, womit fehlende Werte bei Abbruch der Befragung ausgeschlossen wurden.

Für die Analysen fand eine Auszählung der fehlenden Werte bei beiden Erhebungsmethoden und die Überprüfung auf statistische Unterschiede statt (Tabelle 8). Eine Überprüfung auf signifikante Unterschiede in der Verteilung fehlender Werte zwischen den beiden Erhebungsmodi wurde mithilfe des z-Tests im Rahmen einer Kreuztabellierung durchgeführt. Das Signifikanzniveau wurde bei 5 % festgelegt. Die Fallzahlen bei den verschiedenen Fragen unterscheiden sich aufgrund der Vielzahl eingesetzter Fragebögen, Zielgruppen und Filterfragen.

6.1.2 Ergebnisse

Die vergleichenden Analysen zur Item-Nonresponse zwischen den beiden Erhebungsmodi SAQ-

Papier und SAQ-Web zeigen, dass grundsätzlich bei beiden Erhebungsmethoden eine nur geringe Item-Nonresponse zu beobachten ist. Allerdings werden auch Unterschiede zwischen den Erhebungsmodi deutlich (Tabelle 8).

Logisch inkonsistente Angaben treten bei SAQ-Papier deutlich häufiger auf als bei SAQ-Web. Beispiele für logisch inkonsistente Angaben in SAQ-Papier-Befragungen sind unerlaubte Mehrfachangaben innerhalb eines Single-Choice Items,

Tabelle 8
Fehlende Werte für ausgewählte Gesundheitsindikatoren nach Erhebungsmodus

	SAQ-Web		SAQ-Papier		SAQ-Web vs. SAQ-Papier p-Wert
	%	n	%	n	
Allgemeiner Gesundheitszustand					
Subjektive Gesundheit – Elternangabe 0–17 Jahre	0,0	760	0,7	3.176	<0,05
Subjektive Gesundheit – Selbstangabe 11–17 Jahre	0,0	323	1,5	1.017	n.s.
Chronische Krankheit – Elternangabe 0–17 Jahre	0,0	760	0,9	3.176	<0,05
Einschränkung im Vergleich zu anderen – Elternangabe 0–17 Jahre	0,0	760	1,1	3.176	<0,05
Erkrankungen im Kindes- und Jugendalter					
Heuschnupfen Lebenszeitprävalenz (diagnostiziert) – Elternangabe 0–17 Jahre	0,8	760	1,1	3.176	n.s.
Neurodermitis Lebenszeitprävalenz (diagnostiziert) – Elternangabe 0–17 Jahre	0,8	760	1,3	3.176	<0,05
Asthma Lebenszeitprävalenz (diagnostiziert) – Elternangabe 0–17 Jahre	0,9	760	1,3	3.176	<0,05
ADHS Lebenszeitprävalenz (diagnostiziert) – Elternangabe 3–17 Jahre	0,0	645	1,1	2.720	<0,05
Kopfschmerzen – Selbstangabe 11–17 Jahre	0,0	323	5,0	1.017	<0,05
Rückenschmerzen – Selbstangabe 11–17 Jahre	0,3	323	8,2	1.017	<0,05
Psychische Auffälligkeiten und Beeinträchtigungen					
SDQ Impactscore – Elternangabe 3–17 Jahre (keine vollständigen Angaben ¹)	0,0	645	1,8	2.720	<0,05
SDQ Gesamtscore – Elternangabe 3–17 Jahre (keine vollständigen Angaben ¹)	0,3	645	6,2	2.720	<0,05
SDQ Impactscore – Selbstangabe 11–17 Jahre (keine vollständigen Angaben ¹)	0,0	323	0,4	1.017	n.s.
SDQ Gesamtscore – Selbstangabe 11–17 Jahre (keine vollständigen Angaben ¹)	2,1	323	4,7	1.017	n.s.
Adipositas von Kindern, Jugendlichen und deren Müttern					
Körpergröße Kind – Elternangabe 0–10 Jahre	0,6	522	4,3	1.905	<0,05
Körpergewicht Kind – Elternangabe 0–10 Jahre	1,7	522	6,1	1.905	<0,05
Körpergröße Kind – Selbstangabe 11–17 Jahre	0,9	323	1,4	1.017	n.s.
Körpergewicht Kind – Selbstangabe 11–17 Jahre	1,9	323	4,1	1.017	n.s.
Körpergröße Mutter – Elternangabe 0–17 Jahre	1,6	760	2,3	3.176	n.s.
Körpergewicht Mutter – Elternangabe 0–17 Jahre	1,7	760	3,2	3.176	<0,05

Fortsetzung nächste Seite

¹ Keine vollständigen Angaben beinhalten fehlende Werte bei einzelnen Items, weshalb der Gesamtscore nicht berechnet werden konnte. Der Anteil komplett fehlender Angaben ist sowohl beim SDQ-Gesamtscore als auch beim SDQ-Impactscore niedrig (<1,0%).

SAQ=Self-Administered Questionnaire; n.s.=nicht signifikant; ADHS=Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung; SDQ=Strengths and Difficulties Questionnaire

Tabelle 8 Fortsetzung
Fehlende Werte für ausgewählte Gesundheitsindikatoren nach Erhebungsmodus

	SAQ-Web		SAQ-Papier		SAQ-Web vs. SAQ-Papier p-Wert
	%	n	%	n	
Tabakkonsum von Jugendlichen und Eltern					
Rauchstatus Befragte/r – Elternangabe 0–17 Jahre	0,0	760	0,5	3.176	<0,05
Rauchen der Mutter in der Schwangerschaft – Elternangabe 0–17 Jahre	0,1	760	0,7	3.176	n.s.
Rauchen Punktprävalenz – Selbstangabe 11–17 Jahre	1,2	323	0,8	1.017	n.s.
Wasserpfeifenkonsum Lebenszeitprävalenz – Selbstangabe 11–17 Jahre	1,2	323	1,4	1.017	n.s.
Passivrauchbelastung – Selbstangabe 11–17 Jahre	1,2	323	2,2	1.017	n.s.
Alkoholkonsum von Jugendlichen					
Lebenszeitprävalenz Alkohol – Selbstangabe 11–17 Jahre	1,2	323	0,5	1.017	n.s.
Häufigkeit alkoholischer Getränke (AUDIT-C ²) – Selbstangabe 11–17 Jahre	2,0	196 ²	1,9	536 ²	n.s.
Anzahl alkoholischer Getränke (AUDIT-C ²) – Selbstangabe 11–17 Jahre	4,4	158 ²	8,8	456 ²	n.s.
Mehr als 6 Getränke zu einem Zeitpunkt (AUDIT-C ²) – Selbstangabe 11–17 Jahre	2,5	158 ²	2,5	523 ²	n.s.
Körperliche Aktivität und Sport bei Kindern, Jugendlichen und Eltern					
Körperliche Aktivität analog WHO – Elternangabe 3–10 Jahre	0,5	407	2,3	1.449	<0,05
Sport Punktprävalenz Kind – Elternangabe 3–10 Jahre	0,0	407	1,9	1.449	<0,05
Sport Punktprävalenz Elternteil – Elternangabe 0–17 Jahre	0,0	760	1,2	3.176	<0,05
Körperliche Aktivität analog WHO – Selbstangabe 11–17 Jahre	1,5	323	0,3	1.017	n.s.
Sport Punktprävalenz – Selbstangabe 11–17 Jahre	1,2	323	0,7	1.017	n.s.
Inanspruchnahme ärztlicher Leistungen durch Kinder und Jugendliche					
Kinderärztin/Kinderarzt – Elternangabe 0–17 Jahre	0,1	640	0,0	3.176	<0,05
Kieferorthopädische Behandlung – Selbstangabe 11–17 Jahre	1,2	323	3,2	1.017	n.s.

² Bei den Fragen des AUDIT-C wurden in den Papierfragebögen teilweise Angaben gemacht, obwohl diese laut Filterführung des Instruments nicht erforderlich waren. Diese gültigen Angaben wurden nicht auf „filterbedingt fehlend“ gesetzt, weshalb sich die Fallzahlen an dieser Stelle unterscheiden.

Bei der Berechnung der Risikoscores gehen diese gültigen Angaben analog der filterbedingt fehlenden als „nein“ ein.

SAQ = Self-Administered Questionnaire; n.s. = nicht signifikant; AUDIT-C = Alcohol Use Disorders Identification Test; WHO = World Health Organization

Fragen, die aufgrund des Designs übersehen werden, übersehene Filterführungen, unleserliche beziehungsweise missverständliche offene Angaben (Freitexte) und inkonsistente handschriftliche Kommentare zu bestimmten Items, welche einen hohen Arbeitsaufwand bei der Überprüfung und Anpassung nach sich ziehen. Bei SAQ-Web-Fragebögen kann das Risiko für logisch inkonsistente Angaben deutlich verringert werden. Technische Beschränkungen während der Programmierung des Fragebogens erlauben, dass keine Mehrfachangaben bei Single-Choice-Items ausgewählt werden können; ungenaue und unplausible Größen- oder Mengenangaben kön-

nen durch die Definition von Grenzen („ranges“) ausgeschlossen werden. Zudem wird die Filterführung in einem SAQ-Web-Design aufwändig vorab definiert. Teilnehmerinnen und Teilnehmer können somit keine Filterverletzungen begehen. Im Unterschied zur Web-Befragung kann innerhalb der papierbasierten Befragung nur mithilfe des Layouts der Filterführungen und/oder Frage- und Antwortkategorien versucht werden, inkonsistente Antworten zu vermeiden. Da dies nur teilweise gelingt, müssen die Angaben der Teilnehmenden häufig aufwändig geprüft und gegebenenfalls angepasst werden. Die insgesamt sehr niedrige Item-Nonresponse im Online-Modus

kann mit dem sogenannten „Do Answer Check“ erklärt werden (vgl. Kapitel 3.4).

Die Item-Nonresponse bei SAQ-Web liegt meist in einem sehr niedrigen Bereich von unter 2 %, lediglich wenige Indikatoren weisen eine höhere Item-Nonresponse auf. Dies betrifft den SDQ-Gesamtproblemwert, der sich aus insgesamt 25 Einzelitems zusammensetzt. Da hier keine Ersetzung fehlender Werte vorgenommen wird, kann dieser bereits bei einmaliger Nonresponse nicht mehr berechnet werden. Die andere Ausnahme stellen die sensiblen Fragen zum Alkoholkonsum dar, die teilweise ebenfalls eine höhere Prävalenz fehlender Angaben als 2 % aufweisen.

Die Item-Nonresponse der schriftlichen Befragung liegt bei annähernd allen betrachteten Indikatoren höher als bei der Online-Befragung. Besonders deutlich zeigt sich dies bei den Fragen zu Körpergröße und -gewicht von 0- bis 10-jährigen Kindern (Elternangaben) sowie zum Alkoholkonsum von 11- bis 17-jährigen Kindern und Jugendlichen (Selbstangaben). Signifikante Unterschiede gibt es zudem bei allen elternbezogenen Angaben zur gesundheitsbezogenen Lebensqualität der Kinder im Alter von 7 bis 10 Jahren (nicht dargestellt). Die Item-Nonresponse liegt unter 1 % für SAQ-Web- und zwischen 3,4 % bis 4,4 % für SAQ-Papier-Angaben. Bei den Selbstangaben der 11- bis 17-jährigen liegen keine signifikanten Unterschiede bei der gesundheitsbezogenen Lebensqualität vor. Auch bei den Selbstangaben ist der prozentuale Anteil der fehlenden Werte im SAQ-Papier-Modus höher als bei der Online-Befragung.

Bei den Angaben zu Körpergröße und -gewicht der Kinder liegen teilweise inkonsistente und unplausible Angaben vor, die im Rahmen der Datenbereinigung auf fehlende Werte gesetzt wurden. Die Online-Fragebögen wurden so programmiert, dass bei Angaben außerhalb plausibler Wertebereiche eine Rückfrage zu der Angabe gestellt und damit die Möglichkeit zur Korrektur gegeben wurde. Besonders hoch ist die Item-Nonresponse in den schriftlichen Fragebögen bei der 3-Monats-Prävalenz von Kopfschmerzen mit 5,0 % und von Rückenschmerzen mit 8,4 %. Hier lässt sich der hohe Anteil fehlender Angaben mit einem unvorteilhaften Layout der Fragebögen erklären, was für die Hauptstudie von KiGGS Welle 2 geändert wird.

6.2 Prävalenz- und Mittelwertvergleich von Gesundheitsindikatoren

Die Kernindikatoren der KiGGS-Studie wurden sowohl in den Selbst- als auch in den Proxyangaben durch die Eltern (siehe Kapitel 4.4) auf Methodenäquivalenz zwischen den beiden Erhebungsmethoden SAQ-Papier und SAQ-Web geprüft. Dafür wurden alle Teilnehmenden unabhängig vom Erhebungsdesign in die entsprechenden Modegruppen eingeteilt. Es wurden die auf gültigen Werten basierenden Prävalenzen und Mittelwerte ohne Berücksichtigung fehlender Angaben statistisch auf Unterschiedlichkeit getestet. Bei Indikatoren, die aus mehreren Items berechnet werden, fand keine Ersetzung fehlender Werte statt. Eine Ausnahme ist die Berechnung der Scores für die gesundheitsbezogene Lebensqualität, bei der fehlende Werte mit einem vorgegebenen Programm ersetzt wurden [126].

6.2.1 Methodik

Wie bereits in Kapitel 5.3.1 aufgezeigt, unterscheiden sich in beiden Befragungsgruppen (Eltern und 11- bis 17-jährige Kinder und Jugendliche) die beiden Erhebungsmethoden SAQ-Web und SAQ-Papier in mehreren soziodemografischen Merkmalen. Im Allgemeinen sind keine Effekte der Erhebungsmethode auf soziodemografische Angaben zu verzeichnen [18], weshalb diese Unterschiede auch als real verschiedene Eigenschaften und nicht als unterschiedliches Antwortverhalten gewertet werden können. Zur Identifikation möglicher Modeeffekte ist die statistische Kontrolle dieser Merkmale erforderlich [18, 32]. Als methodische Vorgehensweise zur Adjustierung bieten sich unter anderem der Einsatz von Gewichtungungsverfahren, mit denen Merkmalsverteilungen an offizielle Statistiken angeglichen werden [34, 75] sowie multiple Regressionsmodelle [34, 150], Propensity Score Matching [18, 85, 151] oder Strukturgleichungsmodelle [34] an.

Zur Adjustierung soziodemografischer Einflussfaktoren wurden in dieser methodischen Pilotstudie multiple lineare oder logistische Regressionsmodelle berechnet, in die als Kovariaten der Erhebungsmodus und soziodemografische Merkmale aufgenommen wurden. Bei den auf Elternangaben basierenden Gesundheitsindikatoren wurden als Kovariaten Eigenschaften des Kindes (Alter,

Geschlecht, Migrationshintergrund), des befragten Elternteils (Verhältnis zum Kind, Alter, Familienstand) sowie des Haushaltes (höchste Bildung im Haushalt, Haushaltsnettoeinkommen) und des Wohnortes (Gemeindegroße, West-/Ostdeutschland) einbezogen. Bei den Angaben der Kinder und Jugendlichen selbst wurde statt dem Alter und dem Verhältnis des befragten Elternteils zum Kind der höchste vorhandene oder angestrebte Schulabschluss des Kindes mit einbezogen. Aufgrund des hohen Anteils fehlender Werte im Haushaltsnettoeinkommen sowie des höchsten Schulabschlusses des Kindes wurden die fehlenden Angaben als zusätzliche Kategorie in die Analyse aufgenommen. Eine weitere Kategorie stellten nicht ermittelbare Werte aufgrund eines nicht vorhandenen zugehörigen Elternfragebogens bei den Analysen der Angaben der Kinder und Jugendlichen (Familienstand des befragten Elternteils sowie höchste Bildung des Haushaltes und Haushaltsnettoeinkommen).

Mit der Statistiksoftware STATA 14.0 wurden auf der Basis der Regressionsmodelle sogenannte „predictive margins“ berechnet, die eine Schätzung der um die einbezogenen Kovariaten adjustierten Prävalenzen bei der logistischen Regression und des adjustierten Mittelwertes bei der linearen Regression für die beiden Modegruppen

ermöglichen [152, 153]. Diese „predictive margins“ repräsentieren den gewichteten Durchschnitt der gewichteten Wahrscheinlichkeiten eines Gesundheitsindikators in jeder Modegruppe [150]. Die Prüfung auf statistische Unterschiede zwischen den beiden Erhebungsmethoden erfolgte mit den p-Werten der Kovariate „Erhebungsmodus“ im entsprechenden Regressionsmodell mit einem Signifikanzniveau von $p < 0,05$.

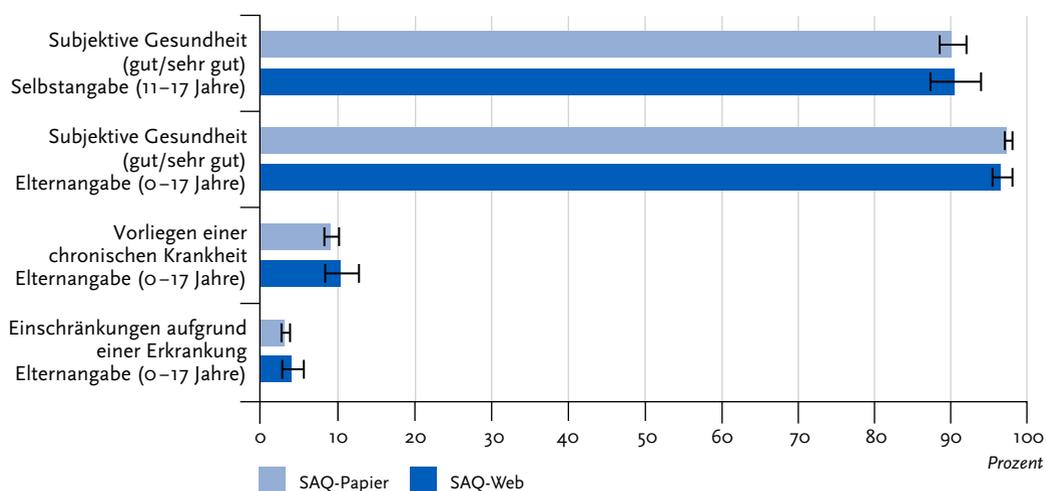
6.2.2 Ergebnisse

Als Fragen zur Einschätzung des allgemeinen Gesundheitszustandes werden die beiden bezüglich der Gesundheit von Kindern und Jugendlichen abgewandelten Fragen aus dem MEHM [154] zur subjektiven Gesundheit und dem Vorhandensein chronischer Krankheiten sowie die Frage zu gesundheitlichen Beeinträchtigungen des Kindes im Vergleich zu anderen Gleichaltrigen des CSHCN-Screeners [111] auf Modeunterschiede analysiert.

Es zeigen sich keine signifikanten Unterschiede in der selbstberichteten Gesundheit aus Eltern- oder Selbstsicht, dem Vorhandensein chronischer Krankheiten oder der gesundheitlichen Einschränkung des Kindes im Vergleich zu Gleichaltrigen (Abbildung 6).

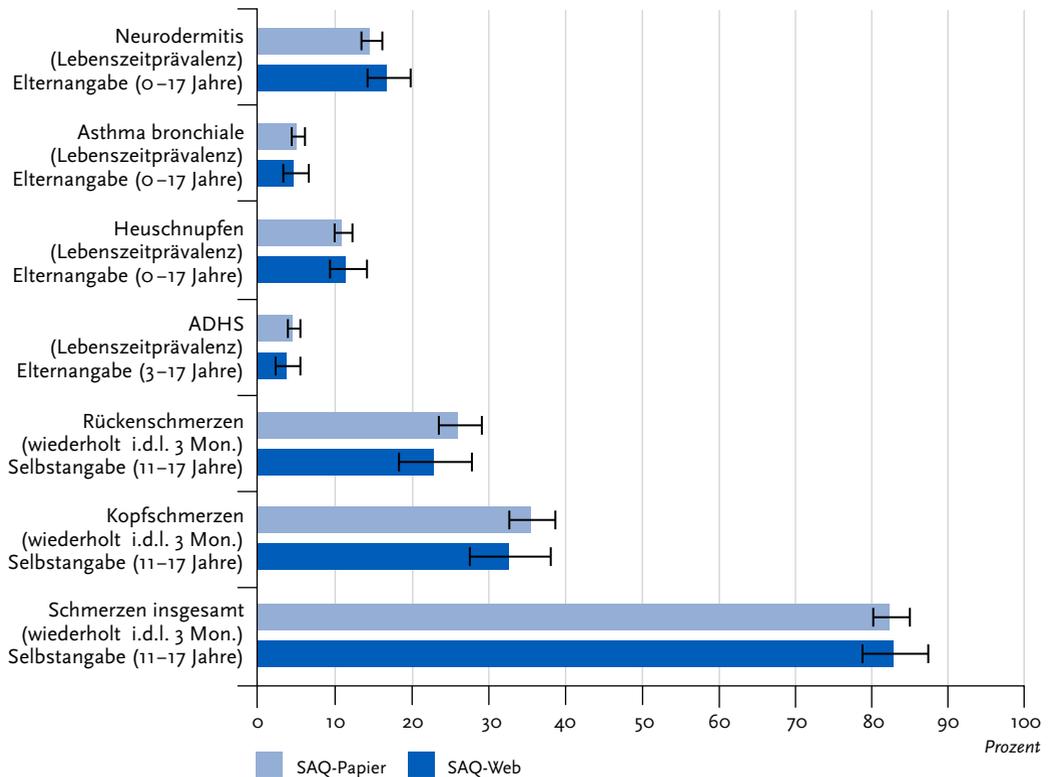
Abbildung 6

Selbsteingeschätzter allgemeiner Gesundheitszustand bei Kindern und Jugendlichen nach Erhebungsmodus (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire

Abbildung 7
Ausgewählte Erkrankungen bei Kindern und Jugendlichen nach Erhebungsmodus
(Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire; ADHS = Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung

Der Großteil anderer Studien findet ebenfalls keinen Unterschied in der subjektiven Gesundheit [73, 150]. In einer methodenvergleichenden Studie wird der Erhebungsmodus abhängig von Internetzugang und Nutzungsverhalten eingeteilt. Personen mit häufiger Internetnutzung und Internetzugang zuhause erhalten den Online-Fragebogen, seltenere Nutzerinnen und Nutzer sowie Personen ohne Internetzugang erhalten den schriftlichen Fragebogen. Zwischen diesen beiden Stichproben ist trotz Kontrolle einer Reihe soziodemografischer Eigenschaften sowie objektiver Krankheiten eine höhere Chance für eine globale gute subjektive Gesundheit in der SAQ-Web-Befragung im Vergleich zur schriftlichen Erhebung vorhanden [72]. Die Autorinnen und Autoren interpretieren den Befund als Ausdruck real unterschiedlicher

Stichprobenmerkmale, die wiederum im Zusammenhang mit dem „digital choice“ (siehe Kapitel 3.2) stehen. Der in anderen Studien gefundene bessere subjektive Gesundheitszustand bei älteren Erwachsenen in der Online-Erhebung [45, 73] wird in den Kontext von empirischen Ergebnissen gestellt, die bei älteren Menschen eine Assoziation zwischen Internetgebrauch und Wohlbefinden beobachten.

Die Prävalenzen wichtiger Erkrankungen im Kindes- und Jugendalter wie die Lebenszeitprävalenz von ADHS, Neurodermitis, Asthma und Heuschnupfen sowie wiederholt auftretender Kopf- oder Rückenschmerzen in den letzten drei Monaten unterscheiden sich nicht statistisch signifikant zwischen den beiden Erhebungsmethoden (Abbildung 7).

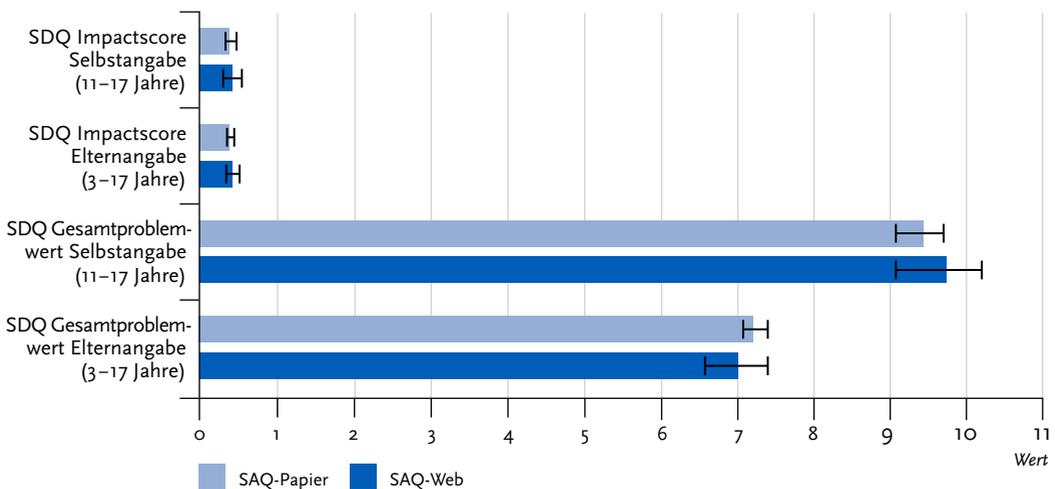
Prävalenzen von Erkrankungen, einzelne Symptome oder Symptomchecklisten werden in anderen Studien ebenfalls als methodenäquivalent zwischen den beiden Selbstausfüllformen SAQ-Web und SAQ-Papier beschrieben [71, 73, 77, 91, 150], diese Befunde werden auch in einer Studie speziell zum selbstberichteten Gesundheitszustand von Jugendlichen bestätigt [155]. Eine norwegische populationsbasierte Gesundheitsstudie findet höhere Asthmaprävalenzen bei Online-Teilnehmenden im Vergleich zu Teilnehmenden der schriftlichen Befragung. Die Autorinnen und Autoren interpretieren diesen Unterschied vor dem Hintergrund gleich hoher Prävalenzen anderer Erkrankungen nicht als Mode-Effekt, sondern als möglichen Nonresponse-Bias [61] und mahnen daher zur Vorsicht, die Web-Alternative nicht ohne Bedenken in eine bisher auf schriftlichen Fragebögen basierende Studie zu integrieren.

Psychische Auffälligkeiten und Beeinträchtigungen bei Kindern und Jugendlichen werden in der KiGGS-Studie mit dem international anerkannten SDQ erhoben, der aus einem Symptom- und einem Impactfragebogen besteht [118, 119]. Für den Altersbereich von 3 bis 17 Jahren liegen Elternangaben vor, Kinder und Jugendliche im Alter von 11 bis 17 Jahre haben den Fragenkatalog zu psychischen Auffälligkeiten selbst ausgefüllt.

Die Prävalenzen der daraus gebildeten Risikogruppen für psychische Auffälligkeiten oder für Beeinträchtigung aufgrund psychischer Probleme (siehe Kapitel 4.4.3) unterscheiden sich weder im Selbst- noch im Proxyurteil durch die Eltern zwischen den beiden analysierten Erhebungsmethoden (Abbildung 8).

Der Befund steht im Einklang mit anderen empirischen Studien, in denen die Messung von Indikatoren psychischer Gesundheit methodenübergreifend verglichen wird. Mehrere Studien finden ebenfalls Methodenäquivalenz zwischen SAQ-Papier und SAQ-Web für standardisierte Instrumente zur Messung von Depressions- und Angstparametern oder Zwangssymptomen sowie bei den Items zur Operationalisierung verschiedener Aspekte psychischer Gesundheit, insbesondere bei Jugendlichen [71, 91, 155–157]. In einem Vergleich verschiedener Instrumente für emotionale und kognitive Dimensionen, Selbstwirksamkeit im Umgang mit negativen Stimmungen und allgemeine Stimmungsregulation werden leichte Unterschiede zwischen den Erhebungsmethoden SAQ-Web und SAQ-Papier gefunden, die allerdings sehr gering und praktisch nicht bedeutsam bewertet werden [158]. Wenige andere Studien machen leichte Einschränkungen, die bei der methodenübergreifenden Messung von Parame-

Abbildung 8
Psychische Auffälligkeiten und Beeinträchtigungen bei Kindern und Jugendlichen nach Erhebungsmodus
(Adjustierte Prävalenzen und 95%-Konfidenzintervalle)

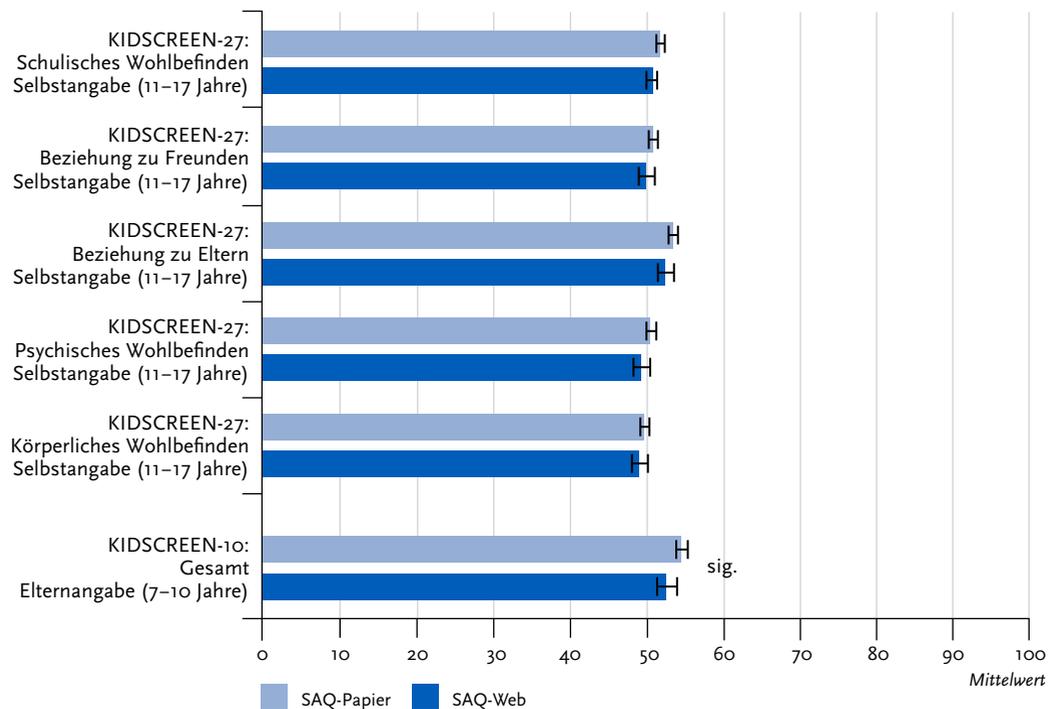


tern psychischer Gesundheit beachtet werden sollen. So zeigen Screeninginstrumente zur Depression online ähnlich valide und reliable Ergebnisse wie die Papierversion [159]. Es wird allerdings eine leicht geringere Depressivität berichtet, woraus höhere Cut-Offs resultieren [159]. Ein weiterer praktischer Aspekt ist, dass es in der Befragung von Indikatoren psychischer Gesundheit wichtig ist, dass Wahlfreiheit für eine Papier- oder Online-Version besteht. Die Begründung liegt darin, dass eine vorhandene Computerversion Befragungsergebnisse in SAQ-Web negativ beeinflussen kann, indem vermehrt negative Gefühle berichtet werden. Dadurch ergeben sich höhere Prävalenzen einer eingeschränkten psychischen Gesundheit wie Depressivität oder Angst, die mit negativen Gefühlen assoziiert sind [160].

Der einzige metrisch analysierte Indikator ist die gesundheitsbezogene Lebensqualität von

Kindern und Jugendlichen. In der selbstberichteten gesundheitsbezogenen Lebensqualität bei den Kindern und Jugendlichen von 11 bis 17 Jahren zeigen sich keine Unterschiede in allen fünf Dimensionen des KIDSCREEN-27 – körperliches und psychisches Wohlbefinden, Beziehung zu den Eltern und zu den Freunden sowie schulisches Wohlbefinden. Der Befund steht im Einklang mit einer methodenvergleichenden Studie zur Messung von Gesundheitsindikatoren bei Jugendlichen, in der ebenfalls keine Unterschiede zwischen den Erhebungsmethoden im körperlichen oder psychischen Wohlbefinden sowie den sozialen Beziehungen gefunden wurde [155]. Die mit dem KIDSCREEN-10 elternberichtete globale gesundheitsbezogene Lebensqualität von Kindern im Alter von 7 bis 10 Jahren zeigt hingegen online eine statistisch niedrigere Lebensqualität (Abbildung 9). Mit Blick auf die relativ geringe

Abbildung 9
Gesundheitsbezogene Lebensqualität bei Kindern und Jugendlichen nach Erhebungsmodus
(Adjustierte Mittelwerte und 95%-Konfidenzintervalle)



Wertebereich 0–40 = schlechte gesundheitsbezogene Lebensqualität (LQ), >40–60 = gute LQ, >60–100 = sehr gute LQ
SAQ = Self-Administered Questionnaire

Mittelwertdifferenz von 2,0 auf einem Range von 0 bis 100 ist es inhaltlich zu diskutieren, ob dieser Unterschied klinisch oder praktisch relevant ist. Die Item-Nonresponse bei der schriftlichen Befragung ist bei den Fragen zur gesundheitlichen Lebensqualität der Kinder deutlich höher als bei der Online-Befragung. Zwar findet im Rahmen der standardisierten Datenaufbereitung des KIDSCREEN-10 [126] eine Bearbeitung fehlender Werte statt, möglicherweise kann eine eventuell vorhandene systematische Item-Nonresponse darüber aber nicht ausgeglichen werden. Eine solche systematische Nonresponse wäre beispielsweise eine erhöhte Wahrscheinlichkeit nicht antwortender Eltern von Kindern mit schlechterer Lebensqualität, die mittels des „Do-Answer-Checks“ bei den Online-Fragebögen doch noch zu gültigen Angaben angeregt werden.

Allerdings zeigt sich beim Vergleich der adjustierten Prävalenzen in die drei kategorialen Einteilungen sehr gute, gute und schlechte gesundheitsbezogene Lebensqualität ein deutlicher Unterschied. 21,3 % der Eltern, die einen schriftlichen Fragebogen nutzen, bewerten die Lebensqualität ihres Kindes als sehr gut und 74,5 % als gut. Im Vergleich dazu schätzen nur 12,9 % der online teilnehmenden Eltern die Lebensqualität ihres Kindes als sehr gut ein und 81,4 % als gut (ohne Abbildung). Eine Literaturrecherche erbrachte keine anderen methodenvergleichenden Studien, in denen das gleiche Instrument in der Proxyberichterstattung durch die Eltern geprüft wird. Eine Vielzahl von Studien vergleicht den SF-36 als ein häufig eingesetztes standardisiertes Instrument bei Erwachsenen und findet Methodenäquivalenz zwischen der webbasierten und der schriftlichen Befragung [71, 91, 102, 161, 162]. Andere Instrumente für die Messung der Lebensqualität von Jugendlichen weisen ebenfalls Methodenäquivalenz auf, lediglich die Einbindung in ein soziales Umfeld mit Gleichaltrigen ist bei online teilnehmenden Jugendlichen besser [155].

Da das Instrument in der Fassung für Selbstangaben durch die Kinder und Jugendlichen in beiden Erhebungsmethoden identisch ausgefüllt wurde und nach empirischer Befundlage eher kein Mode-Effekt in der Messung der gesundheitsbezogenen Lebensqualität zu erwarten ist, erscheint es als unwahrscheinlich, dass der gefundene Unterschied in der elternberichteten Lebensqualität

von 7- bis 10-jährigen Kindern auf Mode-Effekte zurückzuführen ist. Vermutlich geht der Unterschied auf andere Einflussfaktoren zurück, die nicht in die Analysen einbezogen worden sind oder es handelt sich um reale Unterschiede in der Lebensqualität zwischen den beiden Stichproben.

Zur Definition von Adipositas im Kindes- und Jugendalter sowie bei Erwachsenen wird der BMI eingesetzt [129]. Körpergröße und -gewicht wurden von Kindern im Alter von 0 bis 10 Jahren durch die Eltern und im Alter von 11 bis 17 Jahren von Kindern und Jugendlichen selbst im Fragebogen in Zentimetern respektive Kilogramm angegeben. Die Angaben zu Körpergröße und -gewicht von Mutter und Vater erfolgten durch den befragten Elternteil, liegen also entweder als Selbst- oder Proxyangaben vor.

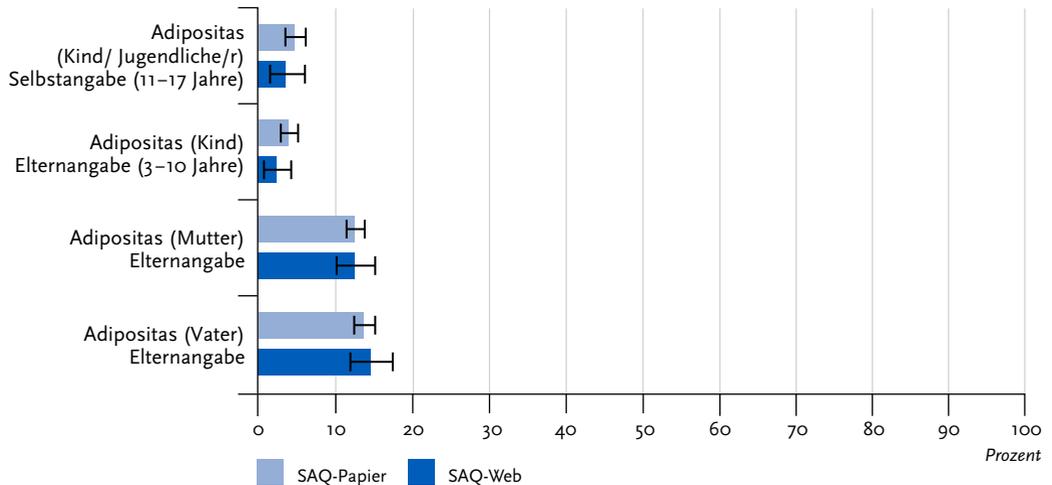
Die auf der Grundlage von Referenzwerten berechneten Adipositasprävalenzen von Kindern, Jugendlichen und deren Eltern unterscheiden sich nicht statistisch signifikant zwischen den beiden Erhebungsmethoden SAQ-Web und SAQ-Papier (Abbildung 10). Die Angaben können in Übereinstimmung mit anderen Studien [77] als methodenäquivalent beurteilt werden. Eine hohe Validität der gemachten Angaben zu Körpergröße und -gewicht in einer schriftlichen oder einer webbasierten Befragung zeigt auch der Vergleich mit dem Goldstandard gemessener Werte, der keine signifikanten Abweichungen ergibt [163].

Die analysierten Indikatoren zum Tabakkonsum von Kindern, Jugendlichen und Eltern zeigen keine statistisch signifikanten Unterschiede zwischen den Erhebungsmethoden (Abbildung 11). Aktueller Tabakkonsum von Kindern, Jugendlichen und deren Eltern, Wasserpfeifenkonsum in den letzten 12 Monaten und Passivrauchbelastung von Kindern und Jugendlichen unterscheiden sich nicht signifikant zwischen den beiden Erhebungsmethoden. Auch sensible Fragen, wie zum Beispiel das mütterliche Rauchen in der Schwangerschaft und Stillzeit, erweisen sich als methodenäquivalent. Diese Befunde sind in Übereinstimmung mit anderen Studienergebnissen für Jugendliche [51, 155, 164] und Erwachsene [77, 78, 150].

Zur Erfassung des Alkoholkonsums von Kindern und Jugendlichen wird zunächst in Form einer Filterfrage die Lebenszeitprävalenz des Alkoholkonsums erhoben. Anschließend wird zur

Abbildung 10

Adipositas bei Kindern, Jugendlichen und deren Eltern nach Erhebungsmodus (Adjustierte Prävalenzen und 95%-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire

Operationalisierung des Risikokonsums und des regelmäßigen Rauschtrinkens das aus drei Fragen bestehende, international etablierte AUDIT-C-Instrument eingesetzt [139].

Es besteht ein statistisch signifikanter Unterschied in der berichteten Lebenszeitprävalenz des Alkoholkonsums zwischen den beiden Erhebungsmethoden (Abbildung 12). In der SAQ-Web-Gruppe berichtet ein um 9,1% höherer Anteil der Kinder und Jugendlichen im Alter von 11 bis 17 Jahren als in der SAQ-Papier-Gruppe, jemals in ihrem Leben Alkohol getrunken zu haben. Die auf den Angaben im AUDIT-C basierenden Prävalenzen zum Risikokonsum und zum regelmäßigen Rauschtrinken unterscheiden sich hingegen nicht signifikant zwischen den Erhebungsmethoden.

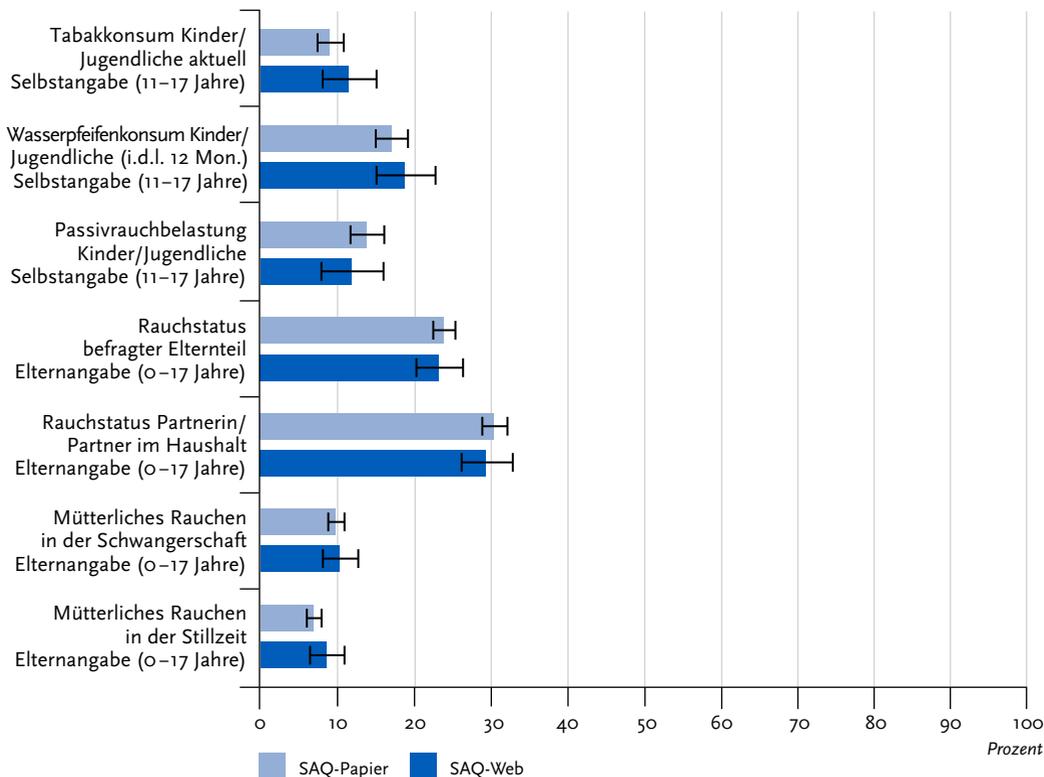
Die Mehrheit anderer Studien berichtet keine Unterschiede in Bezug auf Indikatoren zum Alkoholkonsum bei Erwachsenen [71, 91, 150] sowie bei Jugendlichen oder jungen Erwachsenen [25, 50, 51, 80, 155, 164]. Es gibt aber auch kontroverse Befunde in Studien zu sensiblen Gesundheitsverhalten mit regelmäßigem Rauschtrinken [77] oder allgemeinem Alkoholkonsum [165] bei (jungen) Erwachsenen [25]. Keine Unterschiede fanden sich im sensiblen Gesundheitsverhalten im Allgemeinen bei Collegestudenten [52, 103] oder auch im Konsum illegaler Drogen [166]. Ledig-

lich im Epidemiologischen Suchtsurvey wird bei Erwachsenen ein geringerer Alkoholkonsum in der webbasierten im Vergleich zur schriftlichen Befragung berichtet [58].

Die Ergebnisse einer höheren Lebenszeitprävalenz des Alkoholkonsums bei online antwortenden Kindern und Jugendlichen – ohne signifikanten Unterschied beim regelmäßigen Rauschtrinken oder riskantem Alkoholkonsum – können unter verschiedenen Blickwinkeln interpretiert werden. Möglicherweise basiert dieser Unterschied darauf, dass online teilnehmende Kinder und Jugendliche eher Alkohol ausprobieren als schriftlich teilnehmende Kinder und Jugendliche und sich demzufolge die Stichproben in ihren Eigenschaften unterscheiden. Aber auch die Interpretation als Mode-Effekt bei identischem Alkoholkonsum beider Gruppen ist als Ursache möglich. Eine ehrlichere Berichterstattung im Online-Fragebogen ist vorstellbar bei einer höheren assoziierten Privatheit durch die Jugendlichen, zum Beispiel in dem Sinne, dass ein Online-Fragebogen nicht durch die Eltern kontrolliert werden kann, während ein schriftlicher Fragebogen durch deren Hände gehen könnte. Eine geringere soziale Orientierung analog der Theorie der sozialen Dekontextualisierung durch das Medium Internet [94] ist ebenfalls möglich.

Abbildung 11

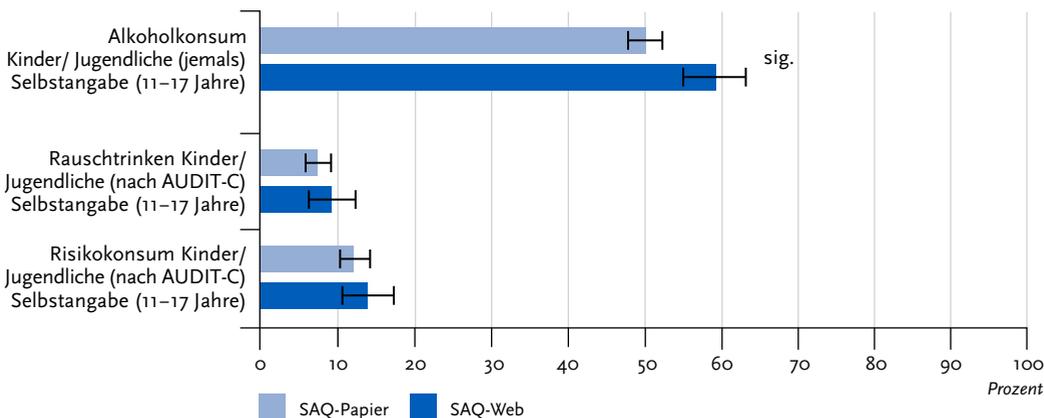
Tabakkonsum von Kindern, Jugendlichen und deren Eltern nach Erhebungsmodus
(Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire

Abbildung 12

Alkoholkonsum von Kindern und Jugendlichen nach Erhebungsmodus (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire; AUDIT-C = Alcohol Use Disorders Identification Test; sig. = signifikant

Körperliche Aktivität und Sport von Kindern, Jugendlichen und deren Eltern erweist sich bei den Kindern und Jugendlichen in allen Fragen als methodenäquivalent zwischen den beiden Erhebungsmethoden (Abbildung 13). Bei den Eltern wird im schriftlichen Fragebogen signifikant weniger von regelmäßigem Sport in den letzten drei Monaten berichtet, was übereinstimmend mit den Ergebnissen einer nationalen Studie bei Erwachsenen ist [150].

Mit Blick auf eine hohe Mediennutzung von Kindern und Jugendlichen von mehr als 2 Stunden Nutzung pro Tag (Abbildung 14) sind ebenfalls keine statistisch signifikanten Unterschiede zwischen den Erhebungsmodi nachzuweisen.

Die Angaben zur Inanspruchnahme ärztlicher Leistungen durch Kinder und Jugendliche sind als methodenäquivalent zu bewerten. Dies gilt

sowohl für die Angaben für einzelne Arztgruppen (hier exemplarisch dargestellt für die Facharztgruppen Kinderheilkunde und Kieferorthopädie, Abbildung 15) als auch für die Inanspruchnahmequote, die aus den Angaben zu allen Arztgruppen berechnet wird. Methodenäquivalenz wird im Kontext der Gesundheitsversorgung für die Impfbeteiligung bei Erwachsenen [150], für das Inanspruchnahmeverhalten bei Jugendlichen [155] und für Merkmale der Versorgungsqualität [63] berichtet.

6.2.3 Fazit zu Mode-Effekten zwischen SAQ-Papier und SAQ-Web

Zusammenfassend kann eine grundsätzlich niedrige Item-Nonresponse bei beiden Erhebungsmethoden festgestellt werden, die bei der schrift-

Abbildung 13

Körperliche Aktivität und Sport bei Kindern, Jugendlichen und Eltern nach Erhebungsmodus (Adjustierte Prävalenzen und 95 %-Konfidenzintervalle)

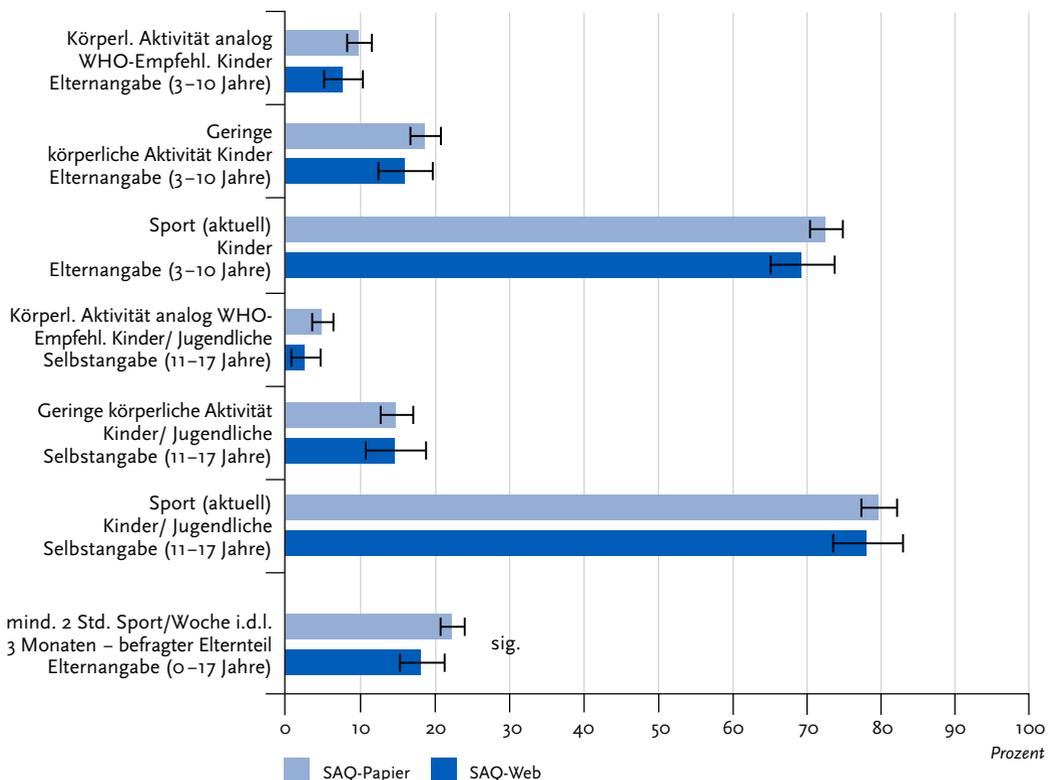
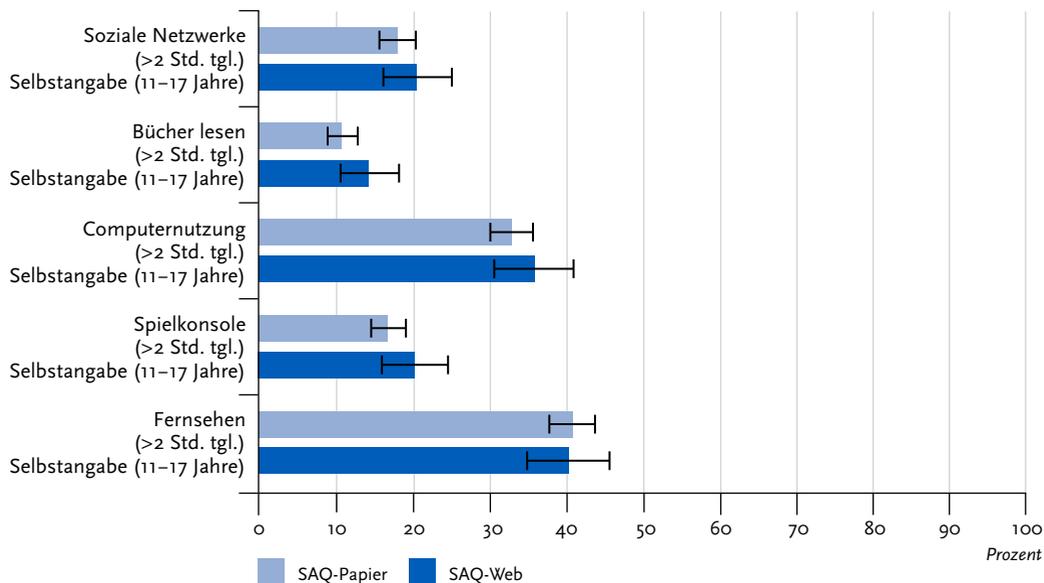


Abbildung 14

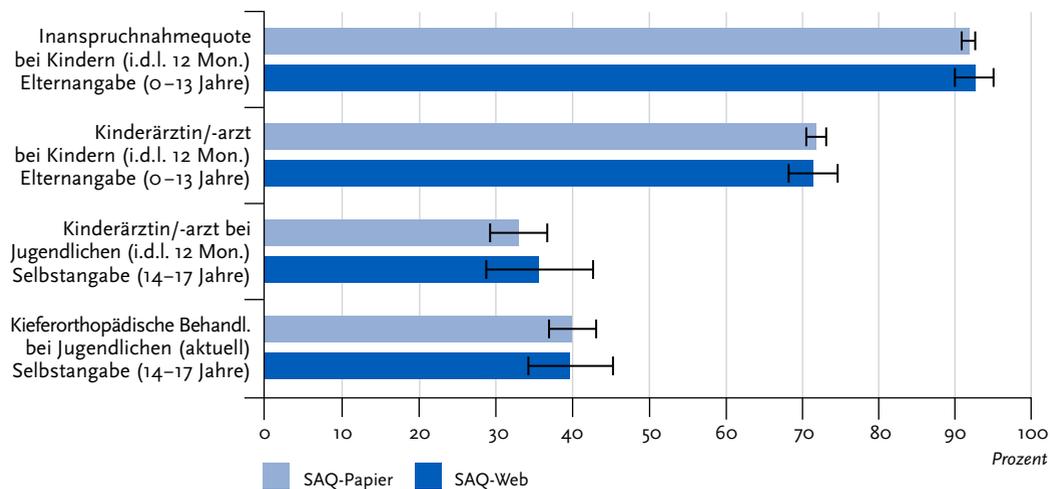
Mediennutzung von Kindern und Jugendlichen nach Erhebungsmodus (Adjustierte Prävalenzen und 95%-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire

Abbildung 15

Inanspruchnahme ärztlicher Leistungen durch Kinder und Jugendliche nach Erhebungsmodus (Adjustierte Prävalenzen und 95%-Konfidenzintervalle)



SAQ = Self-Administered Questionnaire

lichen Befragung höher ist als bei der Online-Befragung. Die Gründe für die unterschiedlich hohe Item-Nonresponse können nicht abschließend geklärt werden. Einen Einfluss können sowohl das leicht unterschiedliche Layout der Fragebögen, die unterschiedliche Filterführung als auch eine mögliche heterogene Zusammensetzung der SAQ-Web- und SAQ-Papier-Teilnehmergruppen sein.

Die meisten der insgesamt 50 analysierten Gesundheitsindikatoren unterscheiden sich nach Adjustierung soziodemografischer Merkmale nicht statistisch signifikant zwischen den beiden Selbstausfüllmethoden SAQ-Web und SAQ-Papier. Signifikante Unterschiede finden sich nur im Lebenszeitkonsum von Alkohol bei Kindern und Jugendlichen, sportlicher Aktivität der Eltern und der elternberichteten gesundheitlichen Lebensqualität von Kindern im Alter von 7 bis 10 Jahren. In der Studie kann nicht geklärt werden, ob sich die gefundenen Unterschiede auf Mode-Effekte oder auf unterschiedliche Eigenschaften der beiden Stichproben zurückführen lassen.

Diese Befunde stehen größtenteils im Einklang mit empirischen Ergebnissen. Die meisten Studien postulieren Methodenäquivalenz für die Messung von Gesundheitsindikatoren bei den beiden Erhebungsmethoden SAQ-Web und SAQ-Papier. Kontroverse Befunde existieren im Kontext von sensiblem Gesundheits- oder Risikoverhalten, das in verschiedenen Studien in der webbasierten Erhebung im Vergleich zur schriftlichen Befragung häufiger berichtet wird. Sensible Fragen sollten in einem Mixed-Mode-Design mit Vorsicht betrachtet werden. Eine konkrete Lösung für die in dieser Studie gefundenen Unterschiede könnte beispielsweise mit Blick auf den Alkoholkonsum sein, dass auf die Berichterstattung der Lebenszeitprävalenz verzichtet wird und nur der Risikokonsum sowie das regelmäßige Rauschtrinken dargestellt werden. Diese beiden Indikatoren weisen Methodenäquivalenz auf und haben eine deutlich höhere Public-Health-Relevanz.

Als abschließendes Fazit der Analysen zu Mode-Effekten kann festgehalten werden, dass eine Mixed-Mode-Erhebung von SAQ-Web und SAQ-Papier unter dem Aspekt eines höheren Messfehlers vertretbar ist, wenn andere Fehlerquellen und/oder die Kosten dadurch verringert werden.

7 Bewertung der Surveyökonomie

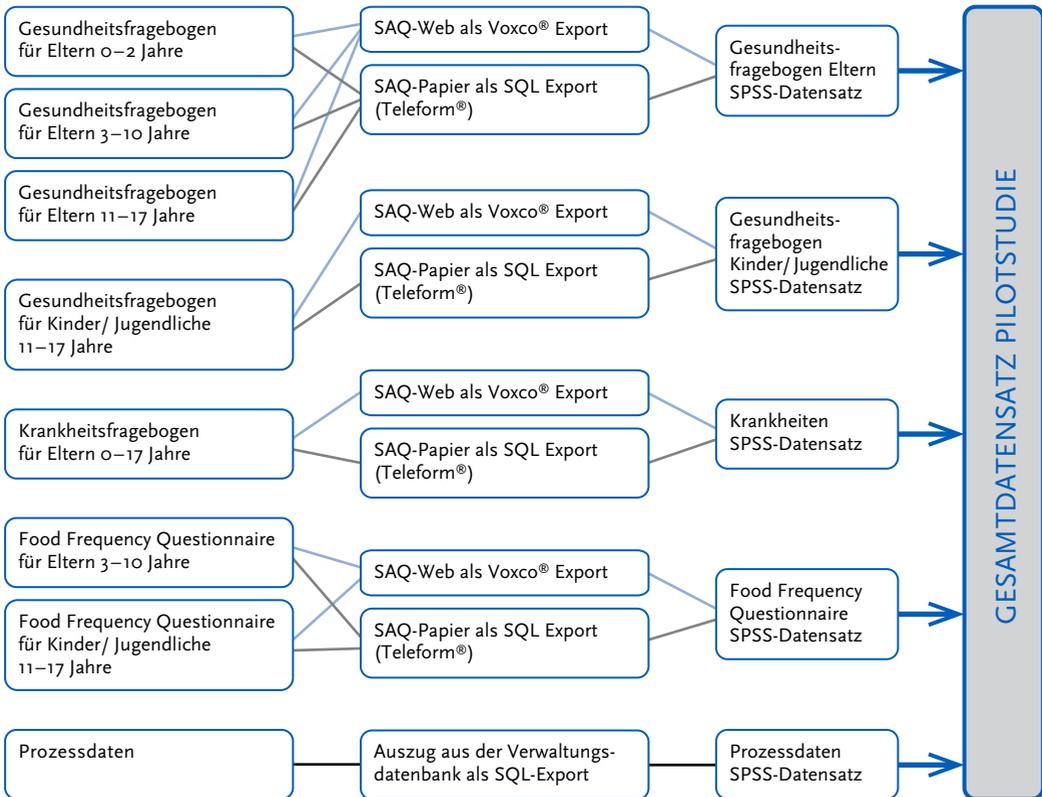
In der methodischen Pilotstudie im Rahmen des Pretests von KiGGS Welle 2 konnten zur Testung eines Mixed-Mode-Designs auf die vorhandene logistische Overhead-Infrastruktur und das Wissen der Mitarbeiterinnen und Mitarbeiter zurückgegriffen werden, das am RKI durch das Labor für Gesundheitsbefragungen (LfG) im Rahmen der Studie GEDA 2.0 aufgebaut worden waren [17].

Die Ergebnisse der methodischen Pilotstudie in KiGGS Welle 2 zeigten beim Vergleich der beiden Erhebungsformen zwei zentrale Vorteile der Web-Befragung auf:

- (1) In der Feldphase gestaltete sich die Datenerhebung mit schriftlichen Fragebögen als zeitlich und personell aufwändiger. Die logistischen Anforderungen durch die Rücklaufverwaltung waren komplexer und es waren zusätzliche Zeit sowie Personal zur Dateneingabe erforderlich. Zudem ist der Aufwand für Lagerung und Archivierung der Fragebögen nur bei SAQ-Papier gegeben.
- (2) Die für die Web-Fragebögen eingesetzte Software hinterlegte für die online eingehenden Angaben der Befragten bereits die Möglichkeit, einen SPSS-Datensatz inklusive der Variablenlabel und Wertebezeichnungen zu exportieren. Hierdurch ergab sich ein sehr geringer Aufwand in der Bearbeitung dieser Daten, die sich überdies durch eine sehr gute Qualität (hohe Plausibilität, stringente Einhaltung der Filterführung, wenige fehlende Angaben) auszeichneten (siehe Kapitel 6.1).

Demgegenüber wurden insbesondere in der Vorbereitungsphase der Pilotstudie umfangreiche Anforderungen sichtbar, die teilweise bei der Studie GEDA 2.0 aufgrund der abweichenden Studiendesigns nicht vorlagen. Dort fand die Befragung mittels eines einheitlichen Fragebogens statt. Dagegen kamen in der methodischen Pilotstudie von KiGGS Welle 2 acht Fragebögen zum Einsatz; in der Hauptphase der KiGGS Welle 2 sind aufgrund der Vielzahl an Zielgruppen letztlich 18 Fragebögen implementiert worden (siehe Kapitel 4.5). Die logistischen Anforderungen waren durch den Einsatz von acht Fragebögen bereits sehr umfangreich:

Abbildung 16
Überblick über die zugrundeliegenden Datensätze



SAQ = Self-Administered Questionnaire; SQL = Structured Query Language; SPSS = Statistiksoftware

- (1) Alle Fragebögen mussten für die schriftliche und für die webbasierte Befragung gesondert programmiert werden. Dies bedeutete einen hohen zeitlichen und personellen Aufwand (identische Umsetzung der Fragebögen, Qualitätssicherung der programmierten Fragebögen). Dadurch erhöhten sich der benötigte zeitliche Vorlauf und die Kosten vor dem Beginn der Datenerhebungen. Von der veränderten zeitlichen Planung war auch die inhaltliche Vorbereitung der Fragebögen betroffen, welche früher abgeschlossen werden musste. Auch minimale inhaltliche Veränderungen würden später die Fehleranfälligkeit in den Fragebögen enorm erhöhen und unverhältnismäßigen Mehraufwand bedeuten.
- (2) Die Berücksichtigung verschiedener Teilnahmemöglichkeiten vervielfältigte die Komplexität der Einladungs- und Erinnerungsprozeduren sowie die Rücklaufverwaltung für die Survey-Geschäftsstelle, welche die Datenerhebungen plant und durchführt beziehungsweise für die Teilnehnergewinnung zuständig ist.
- (3) Ein weiterer zusätzlicher zeitlicher und personeller Aufwand entstand durch die erforderliche Homogenisierung der umfangreichen Daten, die aus unterschiedlichen Quellen stammten (Voxco® und Cardiff® Teleform®) (Abbildung 16). Die Daten werden im Zuge der Speicherung auf verschiedenen Servern mit unterschiedlichen Formaten, Variablen-

labeln und teilweise auch Wertebezeichnungen abgelegt, die nicht direkt miteinander vergleichbar sind. Beispielsweise werden Mehrfachangaben technisch unterschiedlich behandelt und sind nur unter hohem Aufwand zu vereinheitlichen.

Dieser erhebliche Aufwand für ein Mixed-Mode-Design sollte stets in der Planung des Studiendesigns mit einbezogen werden. Es ist eine Kosten-Nutzen-Analyse erforderlich, die für das konkret geplante Studiendesign die möglichen Vorteile eines Mixed-Mode-Designs dem zusätzlichen Aufwand für eine notwendige Doppelstruktur gegenüberstellt.

8 Fazit

Das Ziel der methodischen Pilotstudie in Vorbereitung von KiGGS Welle 2 war zu prüfen, ob durch die Implementierung eines zusätzlichen Web-Befragungsangebots in das bisherige papierbasierte Single-Mode-Design sowohl die Responsequote, die Zusammensetzung der Stichprobe als auch die Surveyökonomie in einem bevölkerungsrepräsentativen Survey zur Gesundheit von Kindern und Jugendlichen verbessert werden kann. Dabei wurden insgesamt 11.140 Kinder und Jugendliche im Alter von 0 bis 17 Jahren sowie deren Eltern zufällig auf eine Kontrollgruppe und verschiedene Mixed-Mode-Erhebungsdesigns verteilt. Bei den Mixed-Mode-Designs wurden das sequenzielle Angebot der Befragungsunterlagen (zuerst SAQ-Web und dann SAQ-Papier), das simultane Angebot der Unterlagen (d. h. das gleichzeitige Zusenden eines schriftlichen Fragebogens und des Online-Zugangscodes) sowie ein Preselect-Design verglichen. Bei letzterem wurden die Teilnehmenden in einem ersten Anschreiben grundsätzlich zur Teilnahme eingeladen und dazu aufgefordert, sich das gewünschte Befragungsmedium (schriftlicher Fragebogen oder Online-Zugangscodes) mittels beiliegenden Antwortformulars anzufordern. In der Kontrollgruppe wurde analog zur KiGGS-Basiserhebung ausschließlich die schriftliche Teilnahme (Single-Mode) angeboten.

Zentrale Ergebnisse zu den Fragestellungen der Pilotstudie KiGGS Welle 2

- (1) Es wird keine höhere Responsequote durch das zusätzliche Angebot eines Online-Fragebogens erreicht.

Mit einem ausschließlichen SAQ-Papier-Angebot (Kontrollgruppe im Single-Mode-Design SAQ-Papier) werden mehr Teilnehmende erreicht als mit einem darüber hinaus unterbreiteten SAQ-Web-Angebot. Eine nur geringfügig geringere Responsequote wurde mit dem simultanen Angebot erreicht, gefolgt vom sequenziellen Angebot. Die niedrigste Responsequote wird im Preselect-Erhebungsdesign erzielt.

Die häufigste Nutzung des Web-Angebots findet im sequenziellen Mixed-Mode-Design statt, in dem der Online-Modus zunächst alternativlos angeboten und erst im zweiten Kontakt (Erinnerungsschreiben zur Studienteilnahme) um die Möglichkeit einer schriftlichen Befragung ergänzt wird. In den Mixed-Mode-Designs mit dem parallelen Angebot der beiden Erhebungsmodi (simultanes Design und Preselect-Design) wird das Online-Angebot im Preselect-Design anteilig häufiger gewählt, jedoch weist letzteres Erhebungsdesign eine insgesamt niedrigere Responsequote als die anderen Erhebungsdesigns auf.

Der Papiermodus kann als präferierte Erhebungsform angesehen werden, weil nur bei fehlender Alternative der Online-Modus etwa so häufig genutzt wird wie ersterer.

- (2) Die Stichprobenszusammensetzung wird in einem Mixed-Mode-Design nicht verbessert.

Das experimentelle Studiendesign ermöglicht es, Unterschiede in der Zusammensetzung der (Teilnehmenden-)Stichproben zwischen der Single-Mode-Kontrollgruppe und den einzelnen Mixed-Mode-Erhebungsdesigns als „nicht zufällig“ zu bewerten. Zwar werden mit den eingesetzten Erhebungsmethoden SAQ-Papier und SAQ-Web unterschiedliche Personengruppen erreicht, dies führt aber in keiner der Mixed-Mode-Erhebungsdesigns zu einer als positiv zu bewertenden Veränderung der Zusammensetzung der Teilnehmergruppe. Systematisch werden keine anderen Personengruppen erreicht als im Single-Mode-Design.

- (3) Mit den beiden Erhebungsmethoden SAQ-Papier und SAQ-Web werden inhaltlich vergleichbare Ergebnisse erreicht.

In der methodischen Pilotstudie werden nach Adjustierung soziodemografischer Merkmale nur wenig statistisch signifikante Unterschiede zwischen den beiden Erhebungsformen SAQ-Papier und SAQ-Web gefunden. Als Fazit der Analysen zu Mode-Effekten kann geschlossen werden, dass eine Mixed-Mode-Erhebung in SAQ-Papier und SAQ-Web unter dem Aspekt eines etwas höheren Messfehlers vertretbar ist.

- (4) Die Datenqualität des Web-Modus ist im Vergleich zur schriftlichen Erhebung höher.

Die Online-Erhebung zeichnet sich durch eine geringere Item-Nonresponse sowie höhere Plausibilität der Daten aus. Dadurch verringert sich der Aufwand für die Qualitätssicherung und Aufbereitung der erhobenen Daten.

- (5) Die Surveyökonomie verbessert sich in diesem Pilotstudiendesign nicht durch das zusätzliche Angebot eines Online-Fragebogens in einem Mixed-Mode-Design.

Der Online-Modus zeigt sich als hoch effektiv bezüglich reduzierter Dateneingabe sowie plausiblerer und konsistenterer Daten. Nur in einem sequenziellen Mixed-Mode-Design, in dem zuerst die Online-Teilnahme und erst später die Teilnahme mittels schriftlicher Befragung angeboten werden, ist der Anteil an Online-Teilnehmenden ausreichend hoch dafür, dass diese Vorteile zum Tragen kommen. Durch die Implementierung eines Online-Fragebogens in eine schriftliche Befragung erhöhen sich der logistische Aufwand in der Survey-Geschäftsstelle sowie der Aufwand für doppelte Programmierung und Datenhomogenisierung. In diesem Pilotstudiendesign übersteigt dieser zusätzliche Aufwand aufgrund der Komplexität der Studie mit unterschiedlichem Befragungsprogramm in Abhängigkeit der Altersgruppen und der Vielzahl eingesetzter Fragebögen die erzielten Vorteile auch in einem sequenziellen Mixed-Mode-Design.

Schlussfolgerung

Bei der Abwägung der Vor- und Nachteile des zusätzlichen Angebots der Online-Befragung kann zusammenfassend festgestellt werden, dass weder bei Eltern noch bei Kindern und Jugendlichen von 11 bis 17 Jahren mit einer Verbesserung der Teilnahmebereitschaft oder der Stichprobensammensetzung zu rechnen ist. Für die jungen Erwachsenen der KiGGS-Kohorte lässt die Pilotstudie keine Rückschlüsse zu. Mit Blick auf mögliche Modeunterschiede ist die Kombination von Papier- und Online-Fragebögen vertretbar. Den surveyökonomischen Vorteilen der Web-Befragung wie einem reduzierten Aufwand der Dateneingabe sowie einer besseren Datenqualität stehen bereits in der Pilotstudie – mit noch geringerem Komplexitätsgrad als der Hauptstudie – erhebliche Nachteile einer Vervielfachung der Verwaltungsprozeduren während der Feldphase sowie bei der Datenaufbereitung und –bereinigung gegenüber. Vor diesem Hintergrund kommt in der KiGGS Welle 2 nur die Befragung mit schriftlichen Fragebögen zum Einsatz.

Literatur

1. Kurth BM (2012) Das RKI-Gesundheitsmonitoring – Was es enthält und wie es genutzt werden kann. *Public Health Forum* 20(3):4.e1-4.e3
2. Kurth BM, Lange C, Kamtsiuris P et al. (2009) Gesundheitsmonitoring am Robert Koch-Institut. Sachstand und Perspektiven. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 52(5):557-570
3. Hölling H, Schlack R, Kamtsiuris P et al. (2012) Die KiGGS-Studie. Bundesweit repräsentative Längs- und Querschnittstudie zur Gesundheit von Kindern und Jugendlichen im Rahmen des Gesundheitsmonitorings am Robert Koch-Institut. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 55(6-7):836-842
4. Lange M, Butschalowsky H, Jentsch F et al. (2014) Die KiGGS-Folgebefragung (KiGGS Welle 1). Studiendurchführung, Stichprobendesign und Response. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):747-761
5. Hölling H, Kamtsiuris P, Lange M et al. (2007) Der Kinder- und Jugendgesundheitsurvey (KiGGS): Studienmanagement und Durchführung der Feldarbeit. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):557-566
6. Kamtsiuris P, Lange M, Schaffrath Rosario A (2007) Der Kinder- und Jugendgesundheitsurvey (KiGGS): Stichprobendesign, Response und Nonresponse-Analyse. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):547-556
7. Kurth BM, Kamtsiuris P, Hölling H et al. (2008) The challenge of comprehensively mapping children's health in a nation-wide health survey: Design of the German KiGGS-Study. *BMC Public Health* 8(1):196
8. Kurth BM, Lange C, Kamtsiuris P et al. (2009) Gesundheitsmonitoring am Robert Koch-Institut. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 52(5):557-570
9. Kurth BM, Kamtsiuris P, Hölling H et al. (2016) Strategien des Robert Koch-Instituts zum Monitoring der Gesundheit von in Deutschland lebenden Kindern und Jugendlichen. *Kinder- und Jugendmedizin* 16(3):176-183
10. Robert Koch-Institut (Hrsg) (2008) Lebensphasenspezifische Gesundheit von Kindern und Jugendlichen in Deutschland. Ergebnisse des Nationalen Kinder- und Jugendgesundheitsurveys (KiGGS). Beiträge zur Gesundheitsberichterstattung des Bundes. RKI, Berlin
11. Bundesministerium für Familie, Senioren, Frauen und Jugend (2009) 13. Kinder- und Jugendbericht. BMFSFJ, Berlin
12. Kurth BM, Bergmann KE, Hölling H et al. (2002) Der bundesweite Kinder- und Jugendgesundheitsurvey. *Gesundheitswesen* 64(S 1):3-11
13. Kurth BM (2007) Der Kinder- und Jugendgesundheitsurvey (KiGGS) Ein Überblick über Planung, Durchführung und Ergebnisse unter Berücksichtigung von Aspekten eines Qualitätsmanagements. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):533-546
14. Lange M, Butschalowsky H, Houben R et al. (2015) Wiederteilnahmebereitschaft, Attrition und Nonresponse-Bias der KiGGS-Kohorte. Jahrestagung der deutschen Gesellschaft für Epidemiologie, Potsdam
15. Hölling H, Schlack R, Kamtsiuris P et al. (2012) Die KiGGS-Studie. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 55(6-7):836-842
16. Robert Koch-Institut (2015) Eckdaten zur „Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland“ – KiGGS Welle 2. https://www.kiggs-studie.de/fileadmin/KiGGS-Dokumente/KiGGS2_Eckdaten.pdf (Stand: 22.08.2017)
17. Robert Koch-Institut (Hrsg) (2015) Pilotstudie zur Durchführung von Mixed-Mode-Gesundheitsbefragungen in der Erwachsenenbevölkerung (Projektstudie GEDA 2.0). Beiträge zur Gesundheitsberichterstattung des Bundes. RKI, Berlin
18. Hox JJ, De Leeuw ED, Zijlman EAO (2015) Measurement equivalence in mixed mode surveys. *Frontiers in Psychology* 6:87
19. Groves RM, Lyberg L (2010) Total Survey Error: Past, Present, and Future. *Public opinion quarterly* 74(5):849-879
20. Biemer PP (2010) Total Survey Error: Design, Implementation and Evaluation. *Public Opinion Quarterly* 74(5):817-848
21. Bowling A (2005) Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health* 27(3):281-291
22. De Leeuw ED (2005) To mix or not to mix data collection modes in surveys. *Journal of Official Statistics* 21(5):233-255
23. Galea S, Tracy M (2007) Participation rates in epidemiologic studies. *Annals of epidemiology* 17(9):643-653
24. Revilla M (2010) Quality in unimode and mixed-mode designs: a multitrait-multimethod approach. *Survey Research Methods* 4(3):151-164
25. McMorris BJ, Petrie RS, Catalano RF et al. (2009) Use of Web and In-Person Survey Modes to Gather Data From Young Adults on Sex and Drug Use: An Evaluation of Cost, Time, and Survey Error Based on a Randomized Mixed-Mode Design. *Evaluation Review* 33(2):138-158
26. Haan M, Ongena Y, Aarts K (2015) The Effects of Mixed-Mode Approach Techniques and Mode Choice on Response Rates of Hard-to-Survey Populations. *Stabilität und Veränderung in Europa*:150-153
27. van Gelder MMHJ, Bretveld RW, Roeleveld N (2010) Web-based Questionnaires: The Future in Epidemiology? *American Journal of Epidemiology* 172(11):1292-1298
28. Voogt RJ, Saris WE (2005) Mixed mode designs: Finding the balance between nonresponse bias and mode effects. *Journal of official statistics* 21(3):367
29. De Leeuw ED, JJ H (2011) Internet surveys as part of a mixed mode design. In: Das M, Ester P, Kaczmirek L (Hrsg) *Social and behavioral research and the internet: Advances in applied methods and research strategies*. Taylor & Francis Group, New York, S. 45-76
30. Roberts C (2007) Mixing modes of data collection in surveys: A methodological review. *NCRM Methods Review Papers*. ESRC National Centre for Research Methods, London
31. De Leeuw ED (2013) Mixed-Mode Surveys and the Internet. *Survey Practice* 3(6)

32. Vannieuwenhuyze J, Loosveldt G, Molenberghs G (2010) A Method for Evaluating Mode Effects in Mixed-mode Surveys. *Public Opinion Quarterly* 74(5):1027-1045
33. Couper MP (2011) The Future of Modes of Data Collection. *Public Opinion Quarterly* 75(5):889-908
34. Jäckle A, Roberts C, Lynn P (2010) Assessing the Effect of Data Collection Mode on Measurement. *International Statistical Review* 78(1):3-20
35. De Leeuw ED, Hox J, Dillman D (2008) Mixed-mode Surveys: When and Why. In: De Leeuw ED, Hox J, Dillman D (Hrsg) *International handbook of survey methodology*. Lawrence Erlbaum Assoc Inc, S. 308–316
36. Dillman DA, Smyth JD, Christian LM (2014) *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, 4th Edition. Wiley, New York
37. Börkan B (2010) The mode effect in mixed-mode surveys mail and web surveys. *Social Science Computer Review* 28(3):371-380
38. De Leeuw ED (2008) Choosing the method of data collection. In: de Leeuw ED, Hox JJ, Dillman DA (Hrsg) *International Handbook of Survey Methodology*. Taylor & Francis Group/Lawrence Erlbaum Associates, New York, S. 113–135
39. Balden W (2004) Multi-mode data collection: Benefits and downsides. In: *Conference of the Great Lakes, Chapter of the Marketing Research Association*, Cancun, Mexico
40. Medway RL, Fulton J (2012) When More Gets You Less: A Meta-Analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates. *Public Opinion Quarterly* 76(4):733-746
41. Millar MM, Dillman DA (2011) Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly* 75(2):249-269
42. Shih T-H, Xitao Fan (2008) Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis. *Field Methods* 20(3):249-271
43. Schouten B, van den Brakel J, Buelens B et al. (2013) Disentangling mode-specific selection and measurement bias in social surveys. *Social Science Research* 42(6):1555-1570
44. Klausch T, Hox J, Schouten B (2015) Selection error in single- and mixed mode surveys of the Dutch general population. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178(4): 945-961
45. Bech M, Kristensen MB (2009) Differential response rates in postal and Web-based surveys in older respondents. *Survey Research Methods* 3(1):1-6
46. Manfreda KL, Bosnjak M, Berzelak J et al. (2008) Web surveys versus other survey modes: A meta-analysis comparing response rates. *Journal of the Market Research Society* 50(1):79
47. Couper MP (2007) Issues of representation in eHealth research (with a focus on web surveys). *Am J Prev Med* 32(5 Suppl):S83–89
48. Radon K, Goldberg M, Becklake M et al. (2002) Low Acceptance of an Internet-Based Online Questionnaire by Young Adults. *Epidemiology* 13(6):748-749
49. Couper MP (2000) Web surveys: A review of issues and approaches. *Public opinion quarterly* 64(4):464-494
50. McCabe SE, Boyd CJ, Couper MP et al. (2002) Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. *Journal of Studies on Alcohol* 63(6):755-761
51. Wyrick DL, Bond L (2011) Reducing Sensitive Survey Response Bias in Research on Adolescents: A Comparison of Web-Based and Paper-and-Pencil Administration. *American Journal of Health Promotion* 25(5):349-352
52. Mi Kyung J (2005) Effects of survey mode, gender, and perceived sensitivity on the quality of data regarding sensitive health behaviors. Thesis, Indiana University, United States
53. Kwak N, Radler B (2002) A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics* 18(2):257-274
54. Bandilla W, Couper MP, Kaczmirek L (2014) The Effectiveness of Mailed Invitations for Web Surveys and the Representativeness of Mixed-Mode versus Internet-only Samples. *Survey Practice* 7(4)
55. Greenlaw C, Brown-Welty S (2009) A comparison of web-based and paper-based survey methods testing assumptions of survey mode and response cost. *Evaluation Review* 33(5):464-480
56. Sax LJ, Gilmartin SK, Bryant AN (2003) Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys. *Research in Higher Education* 44(4):409-432
57. McMahan SR, Iwamoto M, Massoudi MS et al. (2003) Comparison of e-mail, fax, and postal surveys of pediatricians. *Pediatrics* 111(4):e299-e303
58. Kraus L, Pabst A (2010) Studiendesign und Methodik des Epidemiologischen Suchtsurveys 2009. *SUCHT-Zeitschrift für Wissenschaft und Praxis/Journal of Addiction Research and Practice* 56(5):315-326
59. Schneider SJ, Cantor D, Malakhoff L et al. (2005) Telephone, internet, and paper data collection modes for the Census 2000 short form. *Journal of Official Statistics* 21(1):89-101
60. Tancreto JG, Zelenak MF, Davis M et al. (2012) 2011 American Community Survey Internet Tests: Results from First Test in April 2011. American community survey research and evaluation report memorandum series. U.S. Census Bureau, Washington D.C.
61. Brøgger J, Nystad W, Cappelen I et al. (2007) No increase in response rate by adding a web response option to a postal population survey: a randomized trial. *Journal of Medical Internet Research* 9(5)
62. van den Berg MH, Overbeek A, van der Pal HJ et al. (2011) Using Web-Based and Paper-Based Questionnaires for Collecting Data on Fertility Issues Among Female Childhood Cancer Survivors: Differences in Response Characteristics. *Journal of Medical Internet Research* 13(3):e76
63. Zuidgeest M, Hendriks M, Koopman L et al. (2011) A Comparison of a Postal Survey and Mixed-Mode Survey Using a Questionnaire on Patients' Experiences With Breast Care. *Journal of Medical Internet Research* 13(3):e68
64. Turner S, Viera L, Marsh S (2010) Offering a Web Option in a Mail Survey of Young Adults: Impact on Survey Quality. The American Association for Public Opinion Research (AAPOR) 65th Annual Conference, Chicago
65. Hohwü L, Lyshol H, Gissler M et al. (2013) Web-Based Versus Traditional Paper Questionnaires: A Mixed-Mode Survey With a Nordic Perspective. *Journal of Medical Internet Research* 15(8):e173
66. Messer B, Edwards M, Dillman D (2012) Determinants of item nonresponse to web and mail respondents in

- three address-based mixed-mode surveys of the general public. *Survey Practice* 5(2)
67. Griffin DH, Fischer DP, Morgan MT (2001) Testing an Internet response option for the American Community Survey. In: 56th Annual Conference of American Association for Public Opinion Research, Montreal, Quebec, Canada
 68. Brennan M (2005) The effect of a simultaneous mixed-mode (mail and Web) survey on respondent characteristics and survey responses. In: ANZMAC 2005 Conference, S. 5–7
 69. Couper MP, Miller PV (2008) Web survey methods introduction. *Public Opinion Quarterly* 72(5):831–835
 70. Beebe TJ, McAlpine DD, Ziegenfuss JY et al. (2012) Deployment of a Mixed-Mode Data Collection Strategy Does Not Reduce Nonresponse Bias in a General Population Health Survey. *Health Services Research* 47(4):1739–1754
 71. Gwaltney CJ, Shields AL, Shiffman S (2008) Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value in Health* 11(2):322–333
 72. Shim J-M, Shin E, Johnson TP (2013) Self-Rated Health Assessed by Web Versus Mail Modes in a Mixed Mode Survey: The Digital Divide Effect and the Genuine Survey Mode Effect. *Medical Care* 51(9):774–781
 73. de Bernardo DH, Curtis A (2013) Using Online and Paper Surveys: The Effectiveness of Mixed-Mode Methodology for Populations Over 50. *Research on Aging* 35(2):220–240
 74. Dolnicar S, Laesser C, Matus K (2009) Online Versus Paper Format Effects in Tourism Surveys. *Journal of Travel Research* 47(3):295–316
 75. Statistische Ämter des Bundes und der Länder (2010) Ergebnisse des Projekts Q-MED/LFS – Quantifizierung von Methodeneffekten unterschiedlicher Erhebungsinstrumente auf die Datenqualität im Labour Force Survey. Statistisches Bundesamt, Wiesbaden
 76. Smith AB, King M, Butow P et al. (2013) A comparison of data quality and practicality of online versus postal questionnaires in a sample of testicular cancer survivors. *Psycho-Oncology* 22(1):233–237
 77. Link MW, Mokdad AH (2005) Alternative Modes for Health Surveillance Surveys: An Experiment with Web, Mail, and Telephone. *Epidemiology* 16(5):701–704
 78. Callas PW, Solomon LJ, Hughes JR et al. (2010) The Influence of Response Mode on Study Results: Offering Cigarette Smokers a Choice of Postal or Online Completion of a Survey. *Journal of Medical Internet Research* 12(4):e46
 79. Rookey BD, Hanway S, Dillman DA (2008) Does a Probability-Based Household Panel Benefit from Assignment to Postal Response as an Alternative to Internet-Only? *Public Opinion Quarterly* 72(5):962–984
 80. McCabe SE, Diez A, Boyd CJ et al. (2006) Comparing web and mail responses in a mixed mode survey in college alcohol use research. *Addictive behaviors* 31(9):1619–1627
 81. Kaplowitz MD, Hadlock TD, Levine R (2004) A comparison of web and mail survey response rates. *Public opinion quarterly* 68(1):94–101
 82. Fleming M (1997) Evaluating adolescent responses to a computer health information system. *Journal of Adolescent Health* 2(20):166
 83. Decieux JPP (2012) Modeeffekte bei Onlineumfragen. Ein multivariater Methodenvergleich unter Zuhilfenahme eines Propensity Score Matchings. Akademische Verlagsgemeinschaft, München
 84. Lugtig P (2013) Measurement Effects in Mixed-Mode Panel Surveys. NCRM-Conference: Web surveys for the general population: How, why and when? London
 85. Lugtig P, Lensvelt-Mulders GJL, Frerichs R et al. (2011) Estimating nonresponse bias and mode effects in a mixed mode survey. *International Journal of Market Research* 53(5):669–686
 86. Krosnick JA, Narayan S, Smith WR (1996) Satisficing in surveys: Initial evidence. *New Directions for Evaluation* 1996(70):29–44
 87. Krosnick J (2000) The Threat of Satisficing in Surveys: The Shortcuts Respondents Take in Answering Questions. *Survey Methods Newsletter* 20(1):4–8
 88. Möhring W, Schlüt D (2013) Standardisierte Befragung: Grundprinzipien, Einsatz und Anwendung. In: Möhring W, Schlüt D (Hrsg) *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft*. Springer-Verlag, Heidelberg, S. 183–201
 89. Luiten A, Schouten B (2013) Mixed-mode including web: Recent developments at Statistics Netherlands. NCRM-Conference: Web surveys for the general population: How, why and when? London
 90. Dillman DA, Phelps G, Tortora R et al. (2009) Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research* 38(1):1–18
 91. Campbell N, Ali F, Finlay AY et al. (2015) Equivalence of electronic and paper-based patient-reported outcome measures. *Quality of Life Research* 24(8):1949–1961
 92. Tourangeau R, Rasinski K, Jobe JB et al. (1997) Sources of error in a survey on sexual behavior. *Journal of Official Statistics* 13:341–366
 93. Tourangeau R, Smith TW (1996) Asking sensitive questions: the impact of data collection mode, question format and question context. *Public opinion quarterly* 60(2):275–304
 94. Taddicken M (2009) Methodeneffekte von Web-Befragungen: Soziale Erwünschtheit vs. Soziale Entkontextualisierung. In: Weichbold M, Bacher J, Wolf C (Hrsg) *Umfrageforschung*. VS Verlag für Sozialwissenschaften, Wiesbaden, S. 85–104
 95. Wang CC, Liu KS, Cheng CL et al. (2013) Comparison of web-based versus paper-and-pencil administration of a humor survey. *Computers in Human Behavior* 29(3):1007–1011
 96. Klausch T, Hox JJ, Schouten B (2013) Measurement Effects of Survey Mode on the Equivalence of Attitudinal Rating Scale Questions. *Sociological Methods & Research* 42(3):227–263
 97. Gaupp N, Kuhnke R, Schweigard E (2006) Vergleich unterschiedlicher Erhebungsmethoden. Arbeitsbericht im Rahmen der Dokumentationsreihe: Methodische Erträge aus dem „DJI-Übergangspanel“. Deutsches Jugendinstitut, Halle
 98. Coons SJ, Gwaltney CJ, Hays RD et al. (2009) Recommendations on Evidence Needed to Support Measurement Equivalence between Electronic and Paper-Based

- Patient-Reported Outcome (PRO) Measures: ISPOR ePRO Good Research Practices Task Force Report. *Value in Health* 12(4):419-429
99. Thielsch MT, Weltzin S (2009) Online-Befragungen in der Praxis. In: Brandenburg T, Thielsch MT (Hrsg) *Praxis der Wirtschaftspsychologie*. MV Wissenschaft, Münster, S. 69–85
 100. Manfreda KL, Vehovar V (2008) Internet surveys. In: de Leeuw ED, Hox JJ, Dillman DA (Hrsg) *International handbook of survey methodology*. Lawrence Erlbaum Assoc Inc, S. 264–284
 101. Link MW, Mokdad AH (2005) Effects of survey mode on self-reports of adult alcohol consumption: a comparison of mail, web and telephone approaches. *Journal of Studies on Alcohol and Drugs* 66(2):239
 102. Kongsved SM, Basnov M, Holm-Christensen K et al. (2007) Response Rate and Completeness of Questionnaires: A Randomized Study of Internet Versus Paper-and-Pencil Versions. *Journal of Medical Internet Research* 9(3):e25
 103. Pealer LN, Weiler RM, Pigg RM et al. (2001) The Feasibility of a Web-Based Surveillance System to Collect Health Risk Behavior Data from College Students. *Health Education & Behavior* 28(5):547-559
 104. Shieh Y-Y (2003) The Effect of Data Collection Modality on Students' Foreign Language Survey Meeting of the American Association for Public Opinion Research, Nashville
 105. Fleming CM, Bowden M (2009) Web-based surveys as an alternative to traditional mail methods. *Journal of Environmental Management* 90(1):284-292
 106. Smyth JD, Dillman DA, Christian LM et al. (2010) Using the Internet to Survey Small Towns and Communities: Limitations and Possibilities in the Early 21st Century. *American Behavioral Scientist* 53(9):1423-1448
 107. De Bruin A, Picavet HS, Nossikov A (1996) *Health interview surveys: towards harmonization of methods and instruments*. WHO Regional Publications. European Series No. 58. WHO, Copenhagen
 108. Lange M, Kamtsiuris P, Stolzenberg H et al. (2007) Messung soziodemographischer Merkmale im Kinder- und Jugendgesundheitsurvey (KiGGS) und ihre Bedeutung am Beispiel der Einschätzung des allgemeinen Gesundheitszustands. *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):578-589
 109. Lampert T, Müters S, Stolzenberg H et al. (2014) Messung des sozioökonomischen Status in der KiGGS-Studie: Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):762-770
 110. Cox B, Van Oyen H, Cambois E et al. (2009) The reliability of the minimum European health module. *International Journal of Public Health* 54(2):55-60
 111. Bethell CD, Read D, Stein RE et al. (2002) Identifying children with special health care needs: development and evaluation of a short screening instrument. *Ambulatory Pediatrics* 2(1):38-48
 112. Neuhauser H, Poethko-Müller C, KiGGS Study Group (2014) Chronische Erkrankungen und impfpräventable Infektionserkrankungen bei Kindern und Jugendlichen in Deutschland: Ergebnisse der KiGGS-Studie - Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):779-788
 113. Schlack R, Hölling H, Kurth BM et al. (2007) Die Prävalenz der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) bei Kindern und Jugendlichen in Deutschland. Erste Ergebnisse aus dem Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5/6):827-835
 114. Schlack R, Mauz E, Hebebrand J et al. (2014) Hat die Häufigkeit elternberichteter Diagnosen einer Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) in Deutschland zwischen 2003–2006 und 2009–2012 zugenommen? *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):820-829
 115. Schlaud M, Atzpodien K, Thierfelder W (2007) Allergische Erkrankungen. Ergebnisse aus dem Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):701-710
 116. Schmitz R, Thamm M, Ellert U et al. (2014) Verbreitung häufiger Allergien bei Kindern und Jugendlichen in Deutschland: Ergebnisse der KiGGS-Studie - Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):771-778
 117. Ellert U, Neuhauser H, Roth-Isigkeit A (2007) Schmerzen bei Kindern und Jugendlichen in Deutschland: Prävalenz und Inanspruchnahme medizinischer Leistungen. Ergebnisse des Kinder- und Jugendgesundheitsurveys (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):711-717
 118. Goodman R (1997) The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry* 38(5):581-586
 119. Goodman R (1999) The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry* 40(5):791-799
 120. Woerner W, Becker A, Friedrich C et al. (2002) Normierung und Evaluation der deutschen Elternversion des Strengths and Difficulties Questionnaire (SDQ): Ergebnisse einer repräsentativen Felderhebung. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 30(2):105-112
 121. Woerner W, Becker A, Rothenberger A (2004) Normative data and scale properties of the German parent SDQ. *European Child & Adolescent Psychiatry* 13(Suppl 2):I13-10
 122. Hölling H, Erhart M, Ravens-Sieberer U et al. (2007) Verhaltensauffälligkeiten bei Kindern und Jugendlichen. Erste Ergebnisse aus dem Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5/6):784-793
 123. Hölling H, Schlack R, Petermann F et al. (2014) Psychische Auffälligkeiten und psychosoziale Beeinträchtigungen bei Kindern und Jugendlichen im Alter von 3 bis 17 Jahren in Deutschland – Prävalenz und zeitliche Trends zu 2 Erhebungszeitpunkten (2003–2006 und 2009–2012). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):807-819
 124. Erhart M, Ravens-Sieberer U, Dickinson HO et al. (2009) Rasch measurement properties of the KIDSCREEN quality of life instrument in children with cerebral palsy and differential item functioning between children with and without cerebral palsy. *Value Health* 12(5):782-792

125. Ravens-Sieberer U, Erhart M, Rajmil L et al. (2010) Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Qual Life Res* 19(10):1487-1500
126. The KIDSCRREN Group (Hrsg) (2006) The KIDSCREEN questionnaires: Quality of life questionnaires for children and adolescents. Pabst Science Publishers, Lengerich
127. Ellert U, Brettschneider AK, Ravens-Sieberer U et al. (2014) Gesundheitsbezogene Lebensqualität bei Kindern und Jugendlichen in Deutschland: Ergebnisse der KiGGS-Studie - Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):798-806
128. Ravens-Sieberer U, Ellert U, Erhart M (2007) Gesundheitsbezogene Lebensqualität von Kindern und Jugendlichen in Deutschland. Eine Normstichprobe für Deutschland aus dem Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):810-818
129. Dietz WH, Robinson TN (1998) Use of the body mass index (BMI) as a measure of overweight in children and adolescents. *J Pediatr* 132(2):191-193
130. Cole TJ, Bellizzi MC, Flegal KM et al. (2000) Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 320(7244):1240-1243
131. Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter (2012) Leitlinien für Diagnostik, Therapie und Prävention. Verabschiedet auf der Konsensus-Konferenz der AGA am 04.10.2012. <https://de.scribd.com/document/249081404/Adipositas-Leitlinie-KJM> (Stand: 22.08.2017)
132. Kurth BM, Schaffrath Rosario A (2007) Die Verbreitung von Übergewicht und Adipositas bei Kindern und Jugendlichen in Deutschland. Ergebnisse des bundesweiten Kinder- und Jugendgesundheitsurveys (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):736-743
133. Kromeyer-Hauschild K, Wabitsch M, Kunze D et al. (2001) Perzentile für den Body-Mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschrift Kinderheilkunde* 149:807-818
134. World Health Organization (2000) Obesity: preventing and managing the global epidemic. Report of a WHO Consultation (WHO Technical Report Series 894), S. 252
135. Lampert T, Kuntz B, KiGGS Study Group (2014) Tabak- und Alkoholkonsum bei 11- bis 17-jährigen Jugendlichen: Ergebnisse der KiGGS-Studie - Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):830-839
136. Lampert T, Thamm M (2007) Tabak-, Alkohol- und Drogenkonsum von Jugendlichen in Deutschland. Ergebnisse des Kinder- und Jugendgesundheitsurveys (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):600-608
137. Kuntz B, Lampert T, KiGGS Study Group (2015) Wasserpfeifenkonsum (Shisha-Rauchen) bei Jugendlichen in Deutschland Ergebnisse der KiGGS-Studie - Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 58(4):467-473
138. Lampert T, List SM (2010) Gesundheitsrisiko Passivrauchen. Robert Koch-Institut, Berlin. www.rki.de/gbe-kompakt (Stand: 18.06.2010)
139. Bush K, Kivlahan DR, McDonell MB et al. (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 158(16):1789-1795
140. Saunders JB, Aasland OG, Babor TF et al. (1993) Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction* 88(6):791-804
141. Hapke U, von der Lippe E, Gaertner B (2013) Riskanter Alkoholkonsum und Rauschtrinken unter Berücksichtigung von Verletzungen und der Inanspruchnahme alkoholspezifischer medizinischer Beratung: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 56(5-6):809-813
142. Rumpf HJ, Wohler T, Freyer-Adam J et al. (2013) Screening questionnaires for problem drinking in adolescents: performance of AUDIT, AUDIT-C, CRAFFT and POSIT. *European Addiction Research*. 19(3):121-127
143. World Health Organization (2010) Global recommendations on physical activity for health. WHO, Geneva
144. Lampert T, Mensink GBM, Romahn N et al. (2007) Körperlich-sportliche Aktivität von Kindern und Jugendlichen in Deutschland. Ergebnisse des Kinder- und Jugendgesundheitsurveys (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):634-642
145. Manz K, Schlack R, Poethko-Müller C et al. (2014) Körperlich-sportliche Aktivität und Nutzung elektronischer Medien im Kindes- und Jugendalter: Ergebnisse der KiGGS-Studie - Erste Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):840-848
146. Lampert T, Sygusch R, Schlack R (2007) Nutzung elektronischer Medien im Jugendalter. Ergebnisse des Kinder- und Jugendgesundheitsurveys (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):643-652
147. Kamtsiuris P, Bergmann E, Rattay P et al. (2007) Inanspruchnahme medizinischer Leistungen. Ergebnisse des Kinder- und Jugendgesundheitsurveys (KiGGS). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 50(5-6):836-850
148. Rattay P, Starker A, Domanska O et al. (2014) Trends in der Inanspruchnahme ambulant-ärztlicher Leistungen im Kindes- und Jugendalter: Ergebnisse der KiGGS-Studie - Ein Vergleich von Basiserhebung und erster Folgebefragung (KiGGS Welle 1). *Bundesgesundheitsbl - Gesundheitsforsch - Gesundheitsschutz* 57(7):878-891
149. American Association for Public Opinion Research (AAPOR) (2011) Standard definitions: final dispositions of case codes and outcome rates for surveys (revised 2011). AAPOR, Deerfield
150. Hoebel J, von der Lippe E, Lange C et al. (2014) Mode differences in a mixed-mode health interview survey among adults. *Archives of Public Health* 72(1):46

151. Krug G, Kriwy P, Carstensen J (2014) Mixed-Mode Designs bei Erhebungen mit sensitiven Fragen: Einfluss auf das Teilnahme-und Antwortverhalten. *LASER* 84:1-38
152. Graubard BI, Korn EL (1999) Predictive Margins with Survey Data. *Biometrics* 55(2):652-659
153. Chang I-M, Gelman R, Pagano M (1982) Corrected group prognostic curves and summary statistics. *Journal of chronic diseases* 35(8):669-674
154. De Palma E, Crialesi R (2003) Comparative Analysis of Minimum European Health Module and Questions Used in Europe. Italian National Institute of Statistics, Rom
155. Mangunkusumo RT, Moorman PW, Van Den Berg-de Ruyter AE et al. (2005) Internet-administered adolescent health questionnaires compared with a paper version in a randomized study. *Journal of Adolescent Health* 36(1):70.e1-70.e6
156. Austin DW, Carlbring P, Richards JC et al. (2006) Internet Administration of Three Commonly Used Questionnaires in Panic Research: Equivalence to Paper Administration in Australian and Swedish Samples of People With Panic Disorder. *International Journal of Testing* 6(1):25-39
157. Coles ME, Cook LM, Blake TR (2007) Assessing obsessive compulsive symptoms and cognitions on the internet: Evidence for the comparability of paper and Internet administration. *Behaviour Research and Therapy* 45(9):2232-2240
158. Fouladi RT, Mccarthy CJ, Moller N (2002) Paper-and-Pencil Or Online?: Evaluating Mode Effects on Measures of Emotional Functioning and Attachment. *Assessment* 9(2):204-215
159. Donker T, van Straten A, Marks I et al. (2010) Brief self-rated screening for depression on the Internet. *Journal of Affective Disorders* 122(3):253-259
160. Schulenberg S, Yutrzenska B (1999) The equivalence of computerized and paper-and-pencil psychological instruments: Implications for measures of negative affect. *Behavior Research Methods, Instruments, & Computers* 31(2):315-321
161. Basnov M, Kongsved SM, Bech P et al. (2009) Reliability of short form-36 in an Internet-and a pen-and-paper version. *Informatics for Health and Social Care* 34(1):53-58
162. Broering J, Paciorek A, Carroll P et al. (2014) Measurement equivalence using a mixed-mode approach to administer health-related quality of life instruments. *Quality of Life Research* 23(2):495-508
163. West R, Gilseman A, Coste F et al. (2006) The ATTEMPT cohort: a multi-national longitudinal study of predictors, patterns and consequences of smoking cessation; introduction and evaluation of internet recruitment and data collection methods. *Addiction* 101(9):1352-1361
164. Webb PM, Zimet GD, Fortenberry JD et al. (1999) Comparability of a computer-assisted versus written method for collecting health behavior information from adolescent patients. *Journal of Adolescent Health* 24(6):383-388
165. Paperny DM, Aono JY, Lehman RM et al. (1990) Computer-assisted detection and intervention in adolescent high-risk health behaviors. *The Journal of Pediatrics* 116(3):456-462
166. McCabe SE (2004) Comparison of Web and Mail Surveys in Collecting Illicit Drug Use Data: A Randomized Experiment. *Journal of Drug Education* 34(1):61-72

Methodische Studie zur Durchführung von Mixed-Mode-Befragungen zur Gesundheit von Kindern und Jugendlichen (Pilotstudie KiGGS Welle 2)
Robert Koch-Institut, 2017

Herausgeber

Robert Koch-Institut
Nordufer 20
13353 Berlin

Internet: www.rki.de/gbe

Twitter: [@rki_de](https://twitter.com/rki_de)

Redaktion

Stefanie Seeling, Dr. Thomas Ziese
Abteilung für Epidemiologie und Gesundheitsmonitoring

Autorinnen und Autoren

Elvira Mauz, Robert Hoffmann, Robin Houben,
Laura Krause, Antje Gößwald, Panagiotis Kamtsiuris

Satz

Alexander Krönke

Druck

RKI-Hausdruckerei

Bezugsquelle

E-Mail: gbe@rki.de

Tel.: 030-18754-3400

Vorgeschlagene Zitierweise

Robert Koch-Institut (Hrsg) (2017) Methodische Studie zur Durchführung von Mixed-Mode-Befragungen zur Gesundheit von Kindern und Jugendlichen (Pilotstudie KiGGS Welle 2). Gesundheitsberichterstattung des Bundes. Gemeinsam getragen von RKI und Destatis. RKI, Berlin

ISBN: 978-3-89606-236-9

DOI: 10.17886/RKI-GBE-2017-098

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.



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

Die „Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland“ (KiGGS) ist Bestandteil des Gesundheitsmonitorings am Robert Koch-Institut. Die KiGGS-Studie erhebt zum einen regelmäßig repräsentative Daten der heranwachsenden Bevölkerung in Deutschland für Querschnitt- und Trendanalysen und ermöglicht zum anderen längsschnittliche Auswertungen im Rahmen der KiGGS-Kohorte (KiGGS-Basiserhebung 2003–2006, KiGGS Welle 1 2009–2012, KiGGS Welle 2 2014–2017). In der methodischen Pilotstudie in Vorbereitung von KiGGS Welle 2 wurde geprüft, ob mit dem zusätzlichen Angebot eines Online-Fragebogens zur schriftlich-postalischen Befragung die Teilnahmebereitschaft steigt, die Verzerrung der Stichprobe verringert wird und die Vergleichbarkeit der Befragungsergebnisse (Datenqualität) gegeben ist. Verschiedene Mixed-Mode-Designs wurden getestet und mit einer Single-Mode-Kontrollgruppe verglichen, zusätzlich wurden surveyökonomische Aspekte einbezogen. Die Ergebnisse werden hinsichtlich ihrer Konsequenzen für die Hauptstudie (KiGGS Welle 2) beurteilt.