

P. Kamtsiuris  
M. Lange

## Der Pretest des bundesweiten Kinder- und Jugendgesundheits surveys: Stichprobendesign

*The Pilot Study of the National Health Interview and Examination Survey for Children and Adolescents: Sampling Design*

Vorstudien

S107

### Zusammenfassung

Das Konzept für einen bundesweiten Kinder- und Jugendgesundheits survey wurde in einem einjährigen Pretest mit 1630 Kindern und Jugendlichen getestet. Ein wesentlicher methodischer Gesichtspunkt dieser Pilotphase war die Überprüfung verschiedener Varianten der Stichprobenrekrutierung hinsichtlich ihrer Realisierbarkeit. Zur Auswahl standen die Ziehung einer Einwohnermelderegisterstichprobe und die einer Klumpenstichprobe von Schulen bzw. Schulklassen für die Altersgruppe 11 bis 15 Jahre. Im Ergebnis und nach Abwägen der jeweiligen Vor- und Nachteile beider Stichproben-Designs wurde eine durchgängige Einwohnermelderegisterstichprobe empfohlen. Die gewonnenen Erkenntnisse und Erfahrungen werden in der im Jahr 2003 beginnenden Hauptphase des Surveys mit ca. 18 000 Kindern und Jugendlichen umgesetzt.

### Schlüsselwörter

Kinder und Jugendliche · Survey · Pretest · Response · Design-Effekt

### Abstract

The concept of a nation-wide Health Survey for Children and Adolescents has been tested over one year in a pilot study with 1,630 participating children and adolescents. A major methodological aspect of this pilot phase was to compare different variants of study sample recruitment regarding their feasibility. There was a choice between a sample originating from the population registry and a cluster sample of schools and school classes for the age group from 11 to 15 years.

As a result, after weighing the advantages and disadvantages of both sampling designs, a continuous population registry sample has been recommended. The acquired knowledge and experience will be translated into reality in 2003 when the main course of the survey, which will include about 18,000 children and adolescents, will start.

### Key words

Children and adolescents · survey · pilot study · response · design effect

### Einleitung

Im Zuge der Stichprobenplanung für die Hauptphase eines bundesweiten Kinder- und Jugendgesundheits surveys wurden in einem Pretest zwei Varianten der Stichprobenrekrutierung überprüft. Zur Auswahl standen die Ziehung einer Einwohnermelderegisterstichprobe und die einer Klumpenstichprobe von Schulen bzw. Schulklassen für die Altersgruppe 11 bis 15 Jahre.

Bei der Entscheidung, welches der beiden Stichprobendesigns – Einwohnermelderegister oder Schule – in der Hauptphase des Kinder- und Jugendgesundheits surveys für den Altersbereich der 11- bis 15-jährigen Kinder und Jugendlichen zum Einsatz kommen soll, sind verschiedene Kriterien zu berücksichtigen. Diese betreffen die Kosteneffizienz der Varianten, den Informationsbedarf und die Komplexität bei der Stichprobenplanung, die Ausschöpfungsquoten, die Höhe der Unterschiede zwischen Teilnehmer und Nicht-Teilnehmer und damit der Validität der Ergebnisse sowie die Höhe des Stichprobenfehlers und die daraus resultierende Präzision der erhobenen Daten.

### Institutsangaben

Robert-Koch-Institut, Berlin

### Korrespondenzadresse

Panagiotis Kamtsiuris · Robert Koch-Institut · Seestraße 10 · 13353 Berlin

### Bibliografie

Gesundheitswesen 2002; 64 Sonderheft 1: S107–S113 © Georg Thieme Verlag Stuttgart · New York · ISSN 0949-7031

Nachfolgend werden das Stichprobendesign des Pretests sowie ausgewählte Ergebnisse aus dem Vergleich der beiden Stichprobenziehungsvarianten dargestellt. Die Grundgesamtheit (Zielpopulation) des Pretests umfasste alle Kinder und Jugendlichen im Alter zwischen 0 und 18 Jahren, die in den vier ausgesuchten Standorten Berlin-Steglitz, Berlin-Friedrichshain, Neuruppin (Brandenburg) und Wesendorf (Niedersachsen) lebten. Ausgeschlossen waren dabei Kinder und Jugendliche in Einrichtungen wie Krankenhäusern sowie Heil- und Pflegeanstalten.

### Stichprobendesign und Zusammensetzung der Teilnehmergruppe

#### Methode A: Einwohnermelderegisterstichprobe

Methode A folgt den Prinzipien eines geschichteten Random-Designs mit folgenden Auswahlstufen: In jedem der vier Sample Points wurden die Stichprobeneinheiten zuerst nach dem Alter stratifiziert. Über die Einwohnermeldeämter wurden dann aus jeder Schicht (Strata), d. h. für die einzelnen Jahrgänge 0 bis 17, die gleiche Anzahl von Personenadressen zufällig gezogen.

Die Ziehung der Stichprobe erfolgte hierbei nach einem mathematischen Zufallsverfahren (systematische Zufallsauswahl mit Startzahl und vorgegebenem Intervall) aus den Adressdateien der Einwohnermeldeämter. Insgesamt wurden 2479 Kinder und Jugendliche zu einer Teilnahme eingeladen. In den 32 vorgesehenen Untersuchungswochen wurden schließlich 1235 über die Einwohnermelderegister rekrutierte Probanden untersucht.

#### Methode B: Schulen-Stichprobe für den Altersbereich 11 bis 15 Jahre

Methode B folgt den Prinzipien eines geschichteten Random-Cluster-Designs. Das Stichprobendesign ist bei dieser Variante eine Kombination aus einem geschichteten Random-Design und einem Cluster-Design für das Alter 11 bis 15 Jahre mit folgenden Auswahlstufen:

- Die Schulen wurden in den einzelnen Sample Points nach Schultyp geschichtet. Aus jeder Schicht wurde eine Schule zufällig ausgesucht.
- In jeder der gezogenen Schulen wurde aus einer bestimmten Klassenstufe eine Klasse (Cluster) zufällig gezogen. Alle Schüler der ausgewählten Klassen bilden die Bruttostichprobe.
- Da die geplante Anzahl der zu untersuchenden Kinder im Pretest für Variante B (Schulstichprobe) auf maximal 120 pro Sample Point begrenzt war, konnten alle möglichen Kombinationen von Schultyp und Klassenstufe nicht in jedem einzelnen Sample Point realisiert werden. Damit jede mögliche Kombination aus Schultyp und Klassenstufe im Pretest einmal vorkommt, waren diese über alle vier Sample Points zu verteilen. Mithilfe eines einfachen Computerprogramms wurden die möglichen Kombinationen in vier Gruppen eingeteilt und den vier Sample Points zufällig zugeordnet.

Tab. 1 zeigt an, wie der Zusammenhang zwischen herangezogenen Klassenstufen und Verteilung der Brutto-Probanden auf die einzelnen Altersjahrgänge aussieht.

Insgesamt wurden in den Schulen aus einer Bruttostichprobe von 543 Probanden 395 Jugendliche im Alter von 11 bis 17 Jahren untersucht und befragt. Nur 355 dieser Probanden lagen in dem vorgesehenen Altersbereich von 11 bis 15 Jahren.

In Tab. 1 ist zu erkennen, dass durch das Heranziehen ganzer Klassen die Verteilung der realisierten Nettostichprobe über die einzelnen Altersjahrgänge nicht nur von der Ausschöpfung, sondern auch von der altersspezifischen Besetzung der ausgesuchten Klassen zum Zeitpunkt der Untersuchung abhängt. Nur wenn diese altersspezifische Besetzung der Klassen zum Zeitpunkt der Untersuchung in der jeweiligen Schule am Anfang der Hauptphase bekannt wäre, ist auch eine gesicherte Stichprobenplanung im Voraus möglich. Da jedoch am Anfang der Hauptphase für die dreijährige Laufzeit der Feldarbeit die benötigte Information nicht vorhanden sein kann, ist eine Stichprobenplanung nicht ohne weiteres machbar. Die fehlende rechtzeitige Information und die Komplexität der Stichprobenplanung sind ein erheblicher Nachteil der Schulvariante.

Tab. 1 Zusammenhang zwischen Klassenstufen und realisierter Bruttostichprobe

Alter (in Jahren)	Klassenstufe						gesamt
	5	6	7	8	9	10	
11	78%	40%					15,5%
12	20%	58%	15%				18,6%
13	2%	3%	71%	23%			19,9%
14			12%	51%	7%		15,3%
15			1%	20%	76%	21%	21,4%
>15			1%	7%	17%	79%	9,4%
<b>Gesamt-anzahl brutto</b>	<b>40</b>	<b>134</b>	<b>106</b>	<b>123</b>	<b>111</b>	<b>29</b>	<b>543</b>

### Teilnehmerresponse im Pretest

#### Variante A: Einwohnermelderegisterstichprobe

Insgesamt waren 290 von den 2479 eingeladenen Probanden (11,7%) als qualitätsneutrale Ausfälle einzustufen. Für die Einstufung von gezogenen Kindern und Jugendlichen als qualitätsneutrale Ausfälle wurde im Pretest eine ähnlich restriktive Definition wie beim Bundes-Gesundheitssurvey für Erwachsene ausgewählt. Demnach wurden als neutrale Ausfälle solche Fälle aus der Bruttostichprobe herausgenommen, bei denen Folgendes zutraf:

- Proband unbekannt/Adresse falsch
- Wohnung, in der das Kind hauptsächlich lebt, liegt außerhalb des Points
- Proband verstorben
- Verständigung mit Proband bzw. dessen Eltern aufgrund von Sprachproblemen unmöglich
- Proband verzogen bzw. wohnt nicht (mehr) im Haushalt
- Proband ist schon 18 Jahre alt
- Proband ist doppelt gezogen, z. B. auch schon in der Schulvariante.

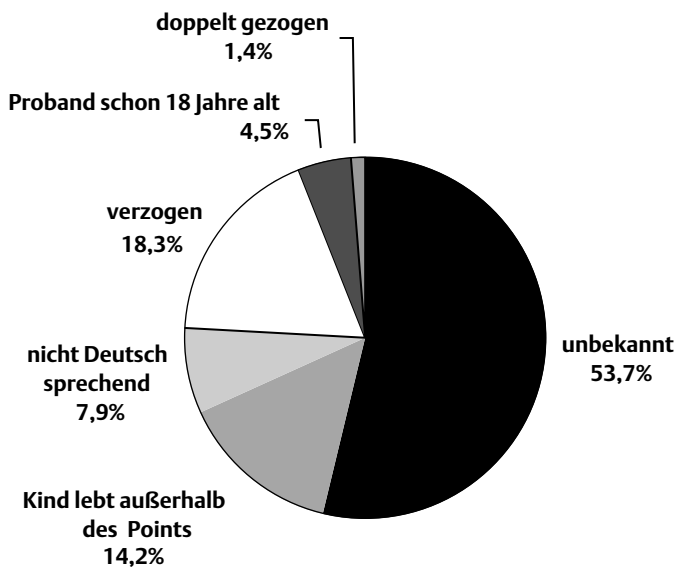


Abb. 1 Zusammensetzung der qualitätsneutralen Ausfälle im Pretest.

Der Anteil der einzelnen Gründe bei der Bestimmung der qualitätsneutralen Ausfälle ist in Abb. 1 abzulesen. Den größten Anteil machten Probanden/Familien aus, die entweder unter der vorliegenden Adresse unbekannt oder zwischenzeitlich verzogen waren. Auch aus anderen Studien ist bekannt, dass die Register der Einwohnermeldeämter nicht den aktuellsten Stand der Wohnbevölkerung wiedergeben. Da die Untersuchungen in den einzelnen Sample Points zum Teil mit einer zeitlichen Verschiebung gegenüber der Ziehung der Stichprobe durchgeführt werden konnten, war dieses Ergebnis zu erwarten. Den drittgrößten Anteil der qualitätsneutralen Ausfälle mit 14,2% (41 Fälle) machten Kinder und Jugendliche aus, die aus verschiedenen Gründen zwar im Untersuchungspoint angemeldet waren, jedoch hauptsächlich außerhalb des Points lebten und deshalb nicht erreichbar waren. Solche Gründe waren unter anderem, dass das Kind bei einem Elternteil angemeldet war, jedoch bei dem anderen Elternteil außerhalb des Points lebte, oder dass das Kind bei den Großeltern im Ausland lebte oder dass das Kind im Internat untergebracht war. Schließlich gab es noch 23 Personen/Familien (ca. 8% der neutralen Ausfälle), die wegen Sprachproblemen an der Studie nicht teilnehmen konnten.

Die Gesamtresponse der Ziehungsvariante Einwohnermelde-register beträgt 57%. Tab. 2 gibt alle wesentlichen Angaben zur Response an. Bei altersspezifischer Betrachtung sind deutliche Unterschiede zu erkennen. Während bei den Säuglingen und den Vorschulkindern (Altersgruppe der 0- bis 6-Jährigen) die Ausschöpfung im Durchschnitt liegt, ist die Response bei den 7- bis 13-Jährigen mit 61% überdurchschnittlich hoch. Die Altersgruppe der 14- bis 17-jährigen Jugendlichen weist demgegenüber mit 50% die niedrigste Responserate aus.

Die Responserate ist bei Mädchen und Jungen etwa gleich hoch. Ein signifikanter geschlechtsspezifischer Unterschied ist nur bei der Altersklasse der 11- bis 13-Jährigen zu verzeichnen.

Insgesamt zeigte sich mit 58% eine höhere Ausschöpfung im Westen (Berlin-Steglitz und Wesendorf/Niedersachsen) als mit 54% im Osten (Berlin-Friedrichshain und Neuruppin/Brandenburg). Die Differenz resultierte in erster Linie aus den Unterschieden in den Altersgruppen 0- bis 2-Jährige und 7- bis 10-Jährige. In den ländlichen Sample Points Neuruppin und Wesendorf war die Ausschöpfung insgesamt mit 59% höher als in den städtischen Erhebungsstandorten Friedrichshain und Steglitz (54%), wobei die höchsten Unterschiede in den Altersgruppen der 3- bis 6-Jährigen bzw. der 7- bis 10-Jährigen zu verzeichnen waren (Tab. 2).

Bei der Stichprobenvariante Einwohnermelderegister waren es insgesamt nach Abzug der qualitätsneutralen Ausfälle 954 Kinder und Jugendliche, die an der Studie nicht teilgenommen haben. Abb. 2 gibt die Häufigkeit an, mit der bestimmte Begründungen für die Nicht-Teilnahme von den Eltern bzw. den Jugendlichen selber oder, für den Fall, dass kein Kontakt mit ihnen zustande kam, vom Untersuchungsteam genannt wurden. Am häufigsten wurden inhaltliche (kein Interesse, vom Sinn und Zweck der Studie nicht überzeugt) und zeitliche Gründe für die Ablehnung genannt. 2,7% der ausgewählten Probanden waren Totalverweigerer, 2,5% sind zum vereinbarten Termin nicht erschienen. Aus gesundheitlichen Gründen hat nur 1,0% an der Studie nicht teilgenommen. Von 11,6% Testpersonen der bereinigten Bruttostichprobe fehlt jedoch eine Begründung für die Nicht-Teilnahme, da diese nie erreicht worden sind.

Inhaltliche Gründe für eine Nicht-Teilnahme wurden vor allem von den 11- bis 17-Jährigen (13,9%), von Kindern und Jugendlichen bzw. deren Familien aus den Erhebungspoints im Westen

Tab. 2 Response – Variante Einwohnermelderegister

gesamt	Altersgruppe	Probanden netto/brutto	Response	Geschlecht		Region		Gemeindegröße	
				weiblich	männlich	Ost	West	Stadt	Land
				n = 1121	n = 1068	n = 1127	n = 1062	n = 1142	n = 1047
	0 bis 2	171/ 312	55%	54%	56%	50%	60%	56%	54%
	3 bis 6	265/ 471	56%	56%	56%	56%	57%	51%	62%
	7 bis 10	290/ 475	61%	61%	62%	54%	68%	57%	65%
	11 bis 13	247/ 402	61%	58%	65%	60%	63%	59%	64%
	14 bis 17	262/ 529	50%	50%	49%	52%	47%	50%	49%
	0 bis 17	1235/2189	56,5%	55,7%	57,2%	54,5%	58,5%	54,2%	58,8%

(11%) sowie von solchen, die in ländlichen Gebieten leben (12,1%), genannt. Zeitliche Gründe kamen mit 9,1% überdurchschnittlich häufig in den Points im Osten sowie in der Winterzeit (9,6%) vor. Gesundheitliche Gründe spielten vor allem in den Wintermonaten eine Rolle (1,6%). Nie erreicht wurden überdurchschnittlich häufig 0- bis 6-Jährige (14,8%), Kinder und Jugendliche im Osten (13,3%), solche, die in städtischen Gebieten wohnen (15,3%), sowie Personen, die für eine Untersuchung in der Sommerzeit ausgewählt worden waren.

**Variante B: Schulstichprobe**

Die Gesamtresponse der Zugangsvariante Schule beträgt bei den 11- bis 15-Jährigen 73% und liegt im Vergleich zur Response der gleichen Altersgruppe in der Variante Einwohnermelderegister um 14% höher. Abb. 3 zeigt, dass die Responsedifferenz von 8% bei den 11-Jährigen über 13 bis 18% bei den 12- bis 14-Jährigen auf 20% bei den 15-Jährigen steigt. Hinsichtlich der Response besitzt die Schulvariante gegenüber der Einwohnermelderegistervariante einen beachtlichen Vorteil.

Tab. 3 gibt alle wesentlichen Angaben zur Response der Schulvariante an. Die Responderate beträgt in den städtischen Sample Points 76% und liegt über der Response in den ländlichen Points (70%). Außerdem ist die Responderate im Westen mit 79% sig-

nifikant höher als im Osten (67%). Dieser Unterschied von 12% resultiert in erster Linie aus der Differenz bei den 13- und 14-Jährigen, wo die Differenz mit 19% am höchsten ist. Auch bei den 11- und 12-Jährigen ist die Differenz mit jeweils 12 bzw. 15% beträchtlich hoch. Allein bei den 15-Jährigen zeigt sich ein umgekehrtes Bild, hier ist die Response im Osten um 7% höher als im

Tab. 3 Response – Variante Schule

Alter (in Jahren)	gesamt Probanden netto/brutto	Response	Region		Gemeindegröße	
			Ost	West	Stadt	Land
11	59/ 82	72%	n = 266 67%	n = 269 79%	n = 304 66%	n = 231 79%
12	72/ 99	73%	64%	79%	77%	67%
13	86/106	81%	71%	90%	89%	73%
14	58/ 83	70%	61%	80%	73%	67%
15	80/116	69%	72%	65%	72%	65%
> 15	40/ 49	82%	71%	88%	88%	50%
gesamt	395/535	74%	68%	80%	77%	70%
gesamt 11-15 Jahre	355/489	73%	67%	79%	76%	70%

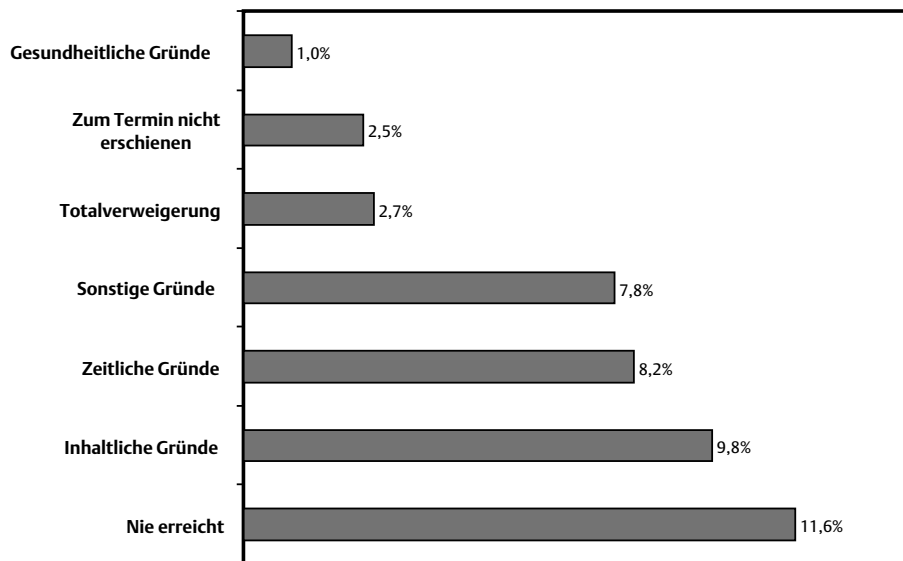


Abb. 2 Gründe für die Nicht-Teilnahme.

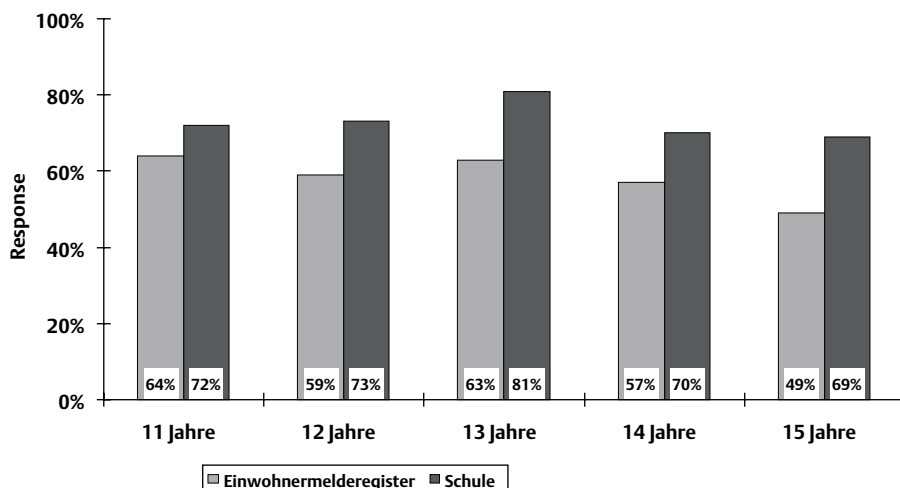


Abb. 3 Vergleich Responderate Einwohnermelderegister vs. Schule.

Westen. Tendenziell ähnliche Unterschiede waren auch in der Einwohnermelderegistervariante zu verzeichnen, die Höhe der Differenzen schwankte jedoch zwischen 3 und 5% und war nicht signifikant.

**Responder vs. Non-Responder und Treffgenauigkeit der Stichprobenvarianten**

Ein weiteres Kriterium für die Beurteilung der beiden Stichprobenvarianten ist das Vorkommen einer Stichprobenverzerrung bzw. eines Non-Response-Bias und dessen Wirkung im Fall einer Nichtberücksichtigung auf die Validität (Treffgenauigkeit) der Ergebnisse. Neben der Ausschöpfungsquote hängt die Stärke eines möglichen Non-Response-Bias davon ab, ob sich Teilnehmer und Nicht-Teilnehmer hinsichtlich ausgewählter Variablen der Erhebung unterscheiden.

Für die Bestimmung von Differenzen zwischen Teilnehmern und Nicht-Teilnehmern sowie für die Berechnung eines möglichen Non-Response-Bias wurde im Pretest ein Kurzfragebogen für die Nicht-Teilnehmer eingesetzt. Im Pretest des Kinder- und Jugendgesundheitsurveys konnten bei der Stichprobenvariante Einwohnermelderegister 469 der 954 Nicht-Teilnehmer nachträglich dazu gewonnen werden, einen Kurzfragebogen zu beantworten. Dieser Anteil entsprach 22% der bereinigten Bruttostichprobe, so dass insgesamt von knapp 80% der Bruttostichprobe Grundinformationen vorliegen. Demgegenüber haben 14% der Personen in der bereinigten Bruttostichprobe (75 von 120 Non-Respondern) in der Stichprobenvariante Schule einen Kurzfragebogen ausgefüllt. Damit ergibt sich ein Anteil von 88%, über den Informationen zumindest über eine begrenzte Anzahl von Grundmerkmalen vorhanden ist.

Auf der Basis des Kurzfragebogens wurden die Randverteilungen ausgewählter Merkmale zwischen Teilnehmern und Nicht-Teilnehmern miteinander verglichen. Tab. 4 zeigt für die Population der 11- bis 15-Jährigen eine Auswahl der herangezogenen

Merkmale und dient auch der Darstellung der Differenzraten dieser Merkmale zwischen Teilnehmern und Nicht-Teilnehmern (hierbei wird die Differenzrate als  $100 \cdot (\text{Teilnehmer} - \text{Nicht-Teilnehmer}) / \text{Teilnehmer}$  definiert).

Während hinsichtlich der geschlechtsdifferenzierten Zusammensetzung die teilnehmende Population der Kinder und Jugendlichen sowohl in der Einwohnermelderegisterstichprobe als auch in der Schulstichprobe kaum Unterschiede zur nichtteilnehmenden Population aufweist, sind in Bezug auf andere wichtige Merkmale (soziale Schicht, Staatsangehörigkeit, gesundheitliche Risikofaktoren) signifikante Unterschiede zu beobachten.

Bei der Einwohnermelderegisterstichprobe ist die Schulbildung der Mutter bei den Non-Respondern als schlechter einzustufen. So haben 86,6% der Non-Responder eine einfache oder mittlere Schulbildung, bei den Respondern sind es nur 70,6%. Die Differenzrate zwischen Respondern und Non-Respondern beträgt -22,7%. Beim Schulabschluss des Vaters zeichnet sich ein ähnliches Bild ab. Hinsichtlich der Berufstätigkeit waren sowohl bei den Müttern als auch bei den Vätern die Vollbeschäftigten stärker bei den Nicht-Teilnehmern repräsentiert. Hinsichtlich der Staatsangehörigkeit, der Frage: „Bei wem wohnt das Kind?“ sowie der Merkmale mit Bezug auf gesundheitliche Risikofaktoren (Rauchen, Gewicht) existieren keine großen Unterschiede zwischen Teilnehmern und Nicht-Teilnehmern.

Bei der Schulstichprobe unterscheiden sich die Teilnehmer hinsichtlich des Schulabschlusses der Mutter kaum von den Nicht-Teilnehmern, der Unterschied bei den Vätern bleibt jedoch bestehen. So haben 75,1% der Väter bei den Teilnehmern einen einfachen oder mittleren Schulabschluss, bei den Nicht-Teilnehmern sind es 82,9%. Die Unterschiede bei der Berufstätigkeit der Eltern sind beim Feldzugang Schule stärker, wobei dies am extremsten bei den Müttern zu beobachten ist. Ist der Anteil der voll berufstätigen Mütter bei den Teilnehmern 41,2%, so sinkt dieser Anteil bei den Nicht-Teilnehmern auf 35,2%. Bei der Frage:

Tab. 4 Randverteilungen ausgewählter Merkmale für Teilnehmer und Nicht-Teilnehmer

Variable	Einwohnermelderegisterstichprobe			Schulstichprobe		
	Teilnehmer (n = 384)	Nicht-Teilnehmer (n = 161)	Differenz-Ratio	Teilnehmer (n = 353)	Nicht-Teilnehmer (n = 62)	Differenz-Ratio
<b>Schulabschluss Mutter</b>						
einfacher + mittlerer	70,6%	86,6%	-22,7%	77,5%	75,0%	3,2%
<b>Schulabschluss Vater</b>						
einfacher + mittlerer	66,4%	80,9%	-21,8%	75,1%	82,9%	-10,4%
<b>Berufstätigkeit Mutter</b>						
voll berufstätig	43,1%	40,3%	6,5%	41,2%	35,2%	14,6%
<b>Berufstätigkeit Vater</b>						
voll berufstätig	82,3%	78,6%	4,5%	83,3%	75,5%	9,4%
<b>Kind wohnt bei</b>						
Eltern	66,8%	68,8%	-3,0%	67,4%	74,1%	-9,9%
<b>Staatsangehörigkeit</b>						
andere	3,2%	3,8%	-18,8%	3,2%	10,2%	-218,8%
<b>Gewicht des Kindes</b>						
Mittelwert	52,74	50,29	4,6%	53,93	58,2	-7,9%

„Bei wem lebt das Kind?“ sind auch stärkere Unterschiede als bei der Einwohnermelderegisterstichprobe zu verzeichnen. Im Gegensatz zur Variante Einwohnermelderegister existiert ein wesentlicher Unterschied beim Anteil der Ausländer in der Teilnehmer- und Nicht-Teilnehmer-Population. Während bei den Non-Respondern die Ausländer 10,2% ausmachen, sind es bei den Teilnehmern nur 3,2%, die keine deutsche Staatsangehörigkeit haben. Hinsichtlich der Merkmale mit Bezug auf gesundheitliche Risikofaktoren sind bei der Schulstichprobe auch stärkere Unterschiede zu beobachten. Das mittlere angegebene Körpergewicht der nicht-teilnehmenden Jugendlichen ist um ca. 4,3 kg höher als das der Teilnehmer. Hier ist jedoch zu beachten, dass das eine das angegebene, subjektiv eingeschätzte Gewicht ist, während das andere gemessen wurde.

### Stichprobenfehler und Präzision der Design-Varianten

Ein weiteres Kriterium für die Beurteilung der beiden Stichprobenvarianten (geschichtetes Random-Design vs. geschichtetes Random-Cluster-Design) ist der Stichprobenfehler. Der Stichprobenfehler lässt sich durch den Standardfehler des Mittelwerts und den Design-Effekt (Deff) bestimmen. Der Standardfehler bzw. standard error (se) des Mittelwerts ist ein Maß für die Präzision, mit der sich der Mittelwert oder der Anteilswert eines Merkmals an der Grundgesamtheit mithilfe der untersuchten Stichprobe schätzen lässt. Je größer die Streuung einer Stichprobe ist, desto größer ist auch der Standardfehler und desto ungenauer ist auch die Mittelwertschätzung. Neben der Streuung variiert der Standardfehler auch mit der Stichprobengröße, d. h., bei steigender Stichprobengröße verringert sich der Standardfehler. Große Stichproben mit geringer Streuung liefern deshalb präzisere Ergebnisse [Kish 1965].

Der Design-Effekt (Deff) beschreibt für komplexe Stichprobendesigns den Einfluss des Ziehungsverfahrens auf die Effizienz der Stichprobe. Der Deff-Wert wird als das Verhältnis der Varianz des eingesetzten komplexen Stichprobendesigns zur Varianz einer einfachen Zufallsauswahl (Random-Design) berechnet. Es handelt sich also um eine Maßzahl für die relative Effizienz eines komplexen Stichprobendesigns im Vergleich zu einem einfachen Random-Design. Ein Deff-Wert von 1 impliziert eine gleiche Effizienz beider Designs. Wenn der Deff-Wert unter 1 liegt, so ist die Varianz des Random-Designs größer als die des komplexen De-

signs, so dass das komplexe Design eine größere Effizienz aufweist. Der umgekehrte Schluss ist für Deff-Werte größer als 1 zu ziehen. Bei einem Deff-Wert von 2 wird beispielsweise beim Random-Design im Vergleich zum komplexen Design die Hälfte der Probanden gebraucht, um die gleiche Präzision der Ergebnisse zu erreichen. Die benötigte Anzahl der Probanden in einem Random-Design wird – um die gleiche Präzision wie in einem komplexen Design zu erreichen – als „Effective n“ bezeichnet [Kalton 1983].

Für den Vergleich der Stichprobendesigns wurde bei beiden Varianten die Gruppe der 11- bis 15-jährigen Probanden herangezogen. Analysiert wurde ein Set von Items (des Eltern- und Kinderfragebogens), aus dem hier beispielhaft zwei Variablen dargestellt werden sollen (Tab. 5). Die Berechnungen wurden mit dem Statistikprogramm SAS durchgeführt, das bei der Kalkulation von Varianz und Standardfehlern das Stichprobendesign berücksichtigt.

Bei den untersuchten Variablen handelt es sich um die Anzahl der Arztbesuche des Kindes in den letzten zwölf Monaten (Angabe der Eltern ergänzt durch Angaben der Kinder/Jugendlichen) sowie das Vorliegen mindestens eines Unfalls des Kindes in den letzten zwölf Monaten (Angabe des Kindes/Jugendlichen). Während sich der Standardfehler bei den Arztbesuchen auf den Mittelwert der Häufigkeiten bezieht, handelt es sich bei den Unfällen um den Anteilswert.

Bei beiden Items ist der Standardfehler des Mittelwerts beim geschichteten Random-Cluster-Design deutlich höher als beim stratifizierten Random-Design (Arztbesuche 0,538 vs. 0,358; Unfälle 0,0358 vs. 0,0268). Da die Items in der Schulvariante eine höhere Varianz aufweisen, sind die Konfidenzintervalle für die Mittelwertschätzung (95%) deutlich breiter als in der Einwohnermelderegisterstichprobe. Damit wird deutlich, dass die Cluster-Stichprobe den Mittelwert der Grundgesamtheit weniger präzise schätzt als die stratifizierte Variante. Wie zudem die Deff-Werte veranschaulichen, ist das Cluster-Design zugleich weniger effektiv als das einfache Random-Design. Denn bei beiden Items ist der Deff-Wert in der Schulvariante deutlich größer als 1. Mit 1,86 bzw. 1,73 liegen die Deff-Werte zudem nahe 2. Das bedeutet, dass die Cluster-Stichprobe im Hauptsurvey fast doppelt so groß sein müsste wie die Random-Stichprobe, um eine ähnlich geringe Varianz zu erreichen.

Tab. 5 Vergleich der Stichprobenvarianten – Standardfehler und Design-Effekt

Merkmal	Zugang	Design	Mittelwert/ Anteil	Standardfehler	Konfidenzintervall		Deff	n im Sample	effective n
					untere Grenze	obere Grenze			
Arztbesuche	EMR*	stratifiziert	5,03	0,358	4,33	5,73	1,00	342	342
		random	5,03	0,359	4,33	5,73		342	
	Schule	cluster	5,39	0,538	4,34	6,44	1,86	322	173
		random	5,39	0,395	4,62	6,16		322	
Unfälle	EMR*	stratifiziert	42,2%	0,0268	36,9%	47,5%	1,00	342	342
		random	42,2%	0,0268	36,9%	47,5%		342	
	Schule	cluster	38,9%	0,0358	31,9%	45,9%	1,73	322	186
		random	38,9%	0,0272	33,6%	44,2%		322	

\* Einwohnermelderegister

### Schlussfolgerungen und Empfehlungen für die Hauptphase

Beim Einwohnermelderegister wurden in den vorgesehenen 32 Untersuchungswochen 1235 Probanden untersucht, d. h. im Durchschnitt 39 Probanden pro Woche. In der Schule nahmen insgesamt 395 Jugendliche in zwölf Wochen teil. Durchschnittlich konnten damit 33 Jugendliche pro Woche untersucht werden. Dementsprechend wird im Fall eines Feldzugangs über die Schule ein größerer Zeitraum für die Untersuchung einer bestimmten Zahl von Probanden benötigt als vergleichsweise beim Einwohnermelderegisterzugang, was zu höheren Untersuchungskosten pro Teilnehmer führt.

Durch das Heranziehen ganzer Klassen hängt die Altersstruktur der realisierten Nettostichprobe nicht nur von der Ausschöpfung, sondern auch von der altersspezifischen Besetzung der ausgesuchten Klassen ab. Die fehlende rechtzeitige Information über die altersspezifische Besetzung der Klassen zum Zeitpunkt der Untersuchung in der jeweiligen Schule und die daraus resultierende Komplexität der Stichprobenplanung sind ein relevanter Nachteil der Schulvariante.

Hinsichtlich der Response besitzt jedoch die Schulvariante gegenüber der Einwohnermelderegistervariante einen beachtlichen Vorteil. Bei einer differenzierten Betrachtung unter Einbeziehung regionaler Aspekte relativiert sich jedoch der Responsevorteil der Schulvariante. Im Fall der Durchführung der Hauptphase auf der Basis einer Schulstichprobe ist eine Verzerrung der Stichprobe zu erwarten, die nur durch eine Gewichtung aufgehoben werden kann. Dieses Problem ist als wesentlicher Nachteil der Schulvariante zu sehen.

Insgesamt lässt sich festhalten, dass in beiden Varianten merkmalspezifische Verzerrungen vorkommen. Während jedoch bei der Einwohnermelderegisterstichprobe diese nur soziodemographische Merkmale betreffen, die z. B. durch den gezielten Einsatz von Incentives und/oder einer Gewichtung der Daten behoben werden können, weist die Schulstichprobe darüber hinaus weitere Merkmalsgruppen mit Verzerrungen auf, wie gesundheitliche Risikofaktoren, die nicht ohne weiteres korrigiert werden können.

Die Analyse der im Pretest erhobenen Daten bestätigte die in der Literatur häufig zu treffende Hypothese, dass Klumpenstichproben gegenüber geschichteten Random-Designs einen Nachteil in Bezug auf die Höhe des Stichprobenfehlers und der damit zusammenhängenden Präzision der Ergebnisse aufweisen.

Aufgrund der oben genannten und der hier nicht diskutierten logistischen Unterschiede [Kurth et al. 2002] der beiden Varianten empfiehlt es sich, in der Hauptphase auch für die Altersgruppe der 11- bis 15-Jährigen ein Einwohnermelderegisterdesign auszuwählen.

### Literatur

- <sup>1</sup> Kalton G. Introduction to Survey Sampling. Thousand Oaks, CA: Sage 1983
- <sup>2</sup> Kurth BM, Bergmann KE, Dippelhofer A et al. Die Gesundheit von Kindern und Jugendlichen in Deutschland. Was wir wissen, was wir nicht wissen, was wir wissen werden. Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz 2002; 45: 852-858
- <sup>3</sup> Kish L. Survey Sampling. New York: John Wiley & Sons Inc 1965