



Gemeinsamer
Bundesausschuss

Beschluss

des Gemeinsamen Bundesausschusses über die Freigabe des
Berichts des IQTIG zur Veröffentlichung:
Konzept für eine onlinebasierte Patientenbefragung im
Rahmen der sektorenübergreifenden Qualitätssicherung nach
§ 136 SGB V

Vom 16. Juni 2022

Der Gemeinsame Bundesausschuss (G-BA) hat in seiner Sitzung am 16. Juni 2022 beschlossen, den Bericht des Instituts für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) „*Konzept für eine onlinebasierte Patientenbefragung im Rahmen der sektorenübergreifenden Qualitätssicherung nach § 136 SGB V*“ gemäß **Anlage** zur Veröffentlichung auf den Internetseiten des IQTIG (www.iqtig.org) freizugeben.

Berlin, den 16. Juni 2022

Gemeinsamer Bundesausschuss
gemäß § 91 SGB V
Der Vorsitzende

Prof. Hecken



Institut für Qualitätssicherung und
Transparenz im Gesundheitswesen

Konzept für eine onlinebasierte Patientenbefragung im Rahmen der sektorenübergreifenden Qualitätssicherung nach § 136 SGB V

Abschlussbericht

Erstellt im Auftrag des
Gemeinsamen Bundesausschusses

Stand: 19. November 2021

Impressum

Thema:

Konzept für eine onlinebasierte Patientenbefragung im Rahmen der sektorenübergreifenden Qualitätssicherung nach § 136 SGB V. Abschlussbericht

Ansprechpartner:

Dr. Gregor Liegl

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

20. November 2020

Datum der Abgabe:

19. November 2021

Herausgeber:

IQTIG – Institut für Qualitätssicherung
und Transparenz im Gesundheitswesen

Katharina-Heinroth-Ufer 1
10787 Berlin

Telefon: (030) 58 58 26-0
Telefax: (030) 58 58 26-999

info@iqtig.org

<https://www.iqtig.org>

Inhaltsverzeichnis

Impressum.....	2
Inhaltsverzeichnis.....	3
Tabellenverzeichnis.....	7
Abbildungsverzeichnis.....	8
Abkürzungsverzeichnis.....	10
Kurzfassung	12
Teil I: Hintergrund und Methoden	18
1 Einleitung.....	19
2 Auftrag und Zielsetzung	21
3 Orientierende Recherchen	23
3.1 Daten zur Internetnutzung in Deutschland.....	23
3.2 Recherchen zur Umsetzungsmöglichkeiten von Onlinebefragungen.....	25
4 Übersicht von systematischen Reviews	26
4.1 Zielsetzung der systematischen Recherche	26
4.2 Präzisierung der Fragestellung.....	26
4.3 Methodisches Vorgehen	27
4.3.1 Einschlusskriterien.....	27
4.4 Informationsquellen.....	28
4.4.1 Bibliographische Datenbanken	28
4.4.2 Handsuche in nicht indextierten Journalen.....	29
4.4.3 Citation searching.....	29
4.5 Selektionsprozess.....	30
4.5.1 Titel- und Abstract-Screening.....	30
4.5.2 Volltext-Screening	30
4.6 Kritische Bewertung.....	30
4.7 Datenextraktion	31
4.8 Überlappung.....	32
4.9 Datensynthese	32
5 Empirische Untersuchung	33
5.1 Zielsetzung und allgemeines Vorgehen	33

5.2	Zielpopulation und verwendete Fragebögen.....	34
5.3	Studiendesign.....	35
5.4	Allokation und Randomisierung.....	36
5.5	Fallzahlkalkulation.....	37
5.6	Rekrutierung der Teilnehmerinnen und Teilnehmer	38
5.7	Feldphase	40
5.8	Statistische Analysen.....	40
Teil II: Ergebnisse		42
6	Nutzung des Internets und von Onlinemedien in Deutschland	43
6.1	Internetnutzung in Deutschland	43
6.2	Nutzung von E-Mail-Diensten	47
6.3	Nutzung von Smartphones.....	49
6.4	Untersuchung des Zusammenhangs zwischen soziodemografischen Merkmalen und der Internetnutzung.....	50
6.5	Multivariable Betrachtung potenzieller Prädiktoren für die Nutzung des Internets	54
7	Möglichkeiten der Umsetzung einer Onlinebefragung mit Patientinnen und Patienten ...	58
7.1	Formen von Onlinebefragungen	58
7.1.1	Single-Mode	59
7.1.2	Mixed-Mode	63
7.2	Design von Onlinebefragungen.....	65
7.2.1	Besonderheiten beim Design von Mixed-Mode-Befragungen	68
7.3	Bewertung der Umsetzung von onlinebasierten Patientenbefragungen im Regelbetrieb.....	69
7.3.1	Datenfluss Patientenbefragung gemäß DeQS-RL.....	69
7.3.2	Kontaktaufnahme bei onlinebasierten Befragungen	71
7.3.2.1	Postalische Kontaktaufnahme	71
7.3.2.2	Elektronische Kontaktaufnahme	72
7.3.3	Umsetzbarkeit von onlinebasierten Single-Mode-Befragungen.....	73
7.3.4	Umsetzbarkeit von Mixed-Mode-Befragungen	73
7.3.5	IT-technische Umsetzung von Onlinebefragungen mittels eines Webportals.....	74
7.3.5.1	Funktionale Anforderungen an das Webportal.....	74
7.3.5.2	Nichtfunktionale Anforderungen an die Software	76

7.3.5.3	Projektanforderungen	77
7.3.5.4	Nutzung eines URL-Kürzungsdienstes	77
7.4	Darstellung der gesetzlichen und datenschutzrechtlichen Rahmenbedingungen ...	78
8	Ergebnisse der Übersicht von systematischen Reviews.....	80
8.1	Eingeschlossene Publikationen	80
8.2	Kritische Bewertung	85
8.3	Effekte des Erhebungsmodus.....	86
8.3.1	Response-Rate.....	86
8.3.2	Antwortverhalten.....	87
8.3.2.1	Vergleiche mit Paper-Pencil-Befragung.....	88
8.3.2.2	Vergleiche elektronischer Erhebungsmodi.....	90
8.4	Zusammenfassung	92
9	Ergebnisse der empirischen Untersuchung.....	94
9.1	Bruttostichprobenbeschreibung	94
9.2	Analysen zum Fragebogenrücklauf und zur Nonresponse	96
9.2.1	Rücklaufquote	96
9.2.2	Beschreibung der Nettostichprobe	98
9.2.3	Moduspräferenzen in den Studienarmen „Mixed-Mode“ und „Push-to-Web“	100
9.3	Item-Nonresponse	102
9.4	Prüfung systematischer Unterschiede bei der Beantwortung der Fragebögen	103
9.5	Zusammenfassung	104
	Teil III: Empfehlungen zur Umsetzung von Onlinebefragungen im Rahmen der gesetzlichen Qualitätssicherung	106
10	Empfehlung zum Einsatz von onlinebasierten Patientenbefragungen im Regelbetrieb ..	107
10.1	Empfehlungen zum Befragungsmodus	107
10.1.1	Bewertung der Umsetzbarkeit und der rechtlichen Rahmenbedingungen unterschiedlicher Befragungsmodi	107
10.1.2	Empfehlung einer Mixed-Mode-Onlinebefragung auf Basis der zu erwartenden Effekte und Konsequenzen für die Datenerhebung.....	108
10.2	Empfehlungen zur Kontaktierung der Befragten	111
10.3	Darstellung des prozessualen Ablaufs bei der Umsetzung einer Mixed-Mode-Befragung mit postalischer Kontaktierung	113

10.3.1 Vorteile und Einsparpotenziale der empfohlenen Umsetzung von onlinebasierten Patientenbefragungen	116
11 Konsequenzen für die Entwicklung und Validierung von onlinebasierten Befragungen ..	117
11.1 Konsequenzen für die Neuentwicklung von Befragungen im Rahmen der gesetzlichen Qualitätssicherung	117
11.2 Überführung bereits bestehender papierbasierter Patientenbefragungen in eine Onlinebefragung	119
Literatur.....	121
Anhang	132

Tabellenverzeichnis

Tabelle 1: PICO-Schema	27
Tabelle 2: Einschlusskriterien.....	28
Tabelle 3: Internetnutzung in Deutschland, Quelle: Destatis (2020) (Destatis 2021b)	43
Tabelle 4: Fragen zur Internetnutzung in den Studien ALLBUS, ESS, VuMA Touchpoints und SHARE.....	45
Tabelle 5: Funktionale Anforderungen an ein Webportal für Onlinebefragungen	74
Tabelle 6: Nicht funktionale Anforderungen an ein Webportal für Onlinebefragungen.....	76
Tabelle 7: Projektanforderungen an ein Webportal für Onlinebefragungen	77
Tabelle 8: Eingeschlossene Artikel	81
Tabelle 9: Verteilung der Publikationen zur Übereinstimmung des Antwortverhaltens	85
Tabelle 10: Überblick über die Bruttostichprobe.....	94
Tabelle 11: Zusammensetzung der bereinigten Bruttostichprobe	95
Tabelle 12: Rücklaufquote nach Einladungsform	97
Tabelle 13: Rücklauf in zeitlichem Bezug zu den Erinnerungsschreiben	97
Tabelle 14: Beschreibung der Nettostichprobe	98
Tabelle 15: Moduspräferenzen in den Studienarmen „Mixed-Mode“ und „Push-to-Web“	101

Abbildungsverzeichnis

Abbildung 1: Design der empirischen Testung	37
Abbildung 2: Angenommene Response-Rate des Vergleichsarmes und benötigte Stichprobengröße	38
Abbildung 3: Anteil Internetnutzerinnen und -nutzer in Deutschland 2018/19, Datenbasis: ESS (ESS Round 9 2021) / ALLBUS (GESIS 2019) / VuMA Touchpoints (VuMa 2021)	46
Abbildung 4: Internetnutzung in Deutschland im Zeitverlauf, Datenbasis: ALLBUS 2012 (GESIS 2013), 2014 (GESIS 2018), 2016 (GESIS 2017) und 2018 (GESIS 2019)	47
Abbildung 5: Häufigkeit der Nutzung von E-Mail-Diensten innerhalb der Gruppe von Internetnutzerinnen und Internetnutzern, Datenbasis: VuMA Touchpoints 2021 (VuMa 2021)	48
Abbildung 6: Smartphonennutzung nach Alter in Deutschland, Quelle: VuMA Touchpoints 2021 (VuMa 2021)	50
Abbildung 7: Internetnutzung nach Bildungsstand und Alter, Datenbasis: ALLBUS 2018 (GESIS 2019) und SHARE Welle 6 (2015) (Börsch-Supan 2019)	51
Abbildung 8: Internetnutzung nach Geschlecht und Alter, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)	52
Abbildung 9: Internetnutzung in Abhängigkeit von gesundheitlichen Einschränkungen im Alltag, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)	53
Abbildung 10: Internetnutzung in Abhängigkeit von der Summe der diagnostizierten Erkrankungen, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)	54
Abbildung 11: Binomial-logistische Regression zu Internetnutzung in den letzten sieben Tagen mit Einschränkungen im Alltag als Indikator für den Gesundheitsstatus, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)	55
Abbildung 12: Binomial-logistische Regression zu Internetnutzung in den letzten sieben Tagen mit Summe der diagnostizierten Erkrankungen als Indikator für den Gesundheitsstatus, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)	56
Abbildung 13: Binomial-logistische Regression zu Internetnutzung in den letzten sieben Tagen mit einzelnen Erkrankungen bzw. Erkrankungsgruppen als Indikatoren für den Gesundheitsstatus, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)	57
Abbildung 14: Datenfluss Patientenbefragung gemäß DeQS-RL	69
Abbildung 15: Flussdiagramm der Recherche nach systematischen Reviews	80
Abbildung 16: Modell zum Zusammenhang zwischen dem Erhebungsmodus und der Anzahl von Missings unter Kontrolle von Alter, Geschlecht, Anzahl der Komorbiditäten und Bildung	103
Abbildung 17: Einfluss des Modus auf die Personen-Scores, die auf Basis der zugrunde liegenden Items von 13 Qualitätsindikatoren der Patientenbefragung QS NET (Bereich Dialyse) berechnet wurden	104
Abbildung 18: Prozessualer Ablauf bei der Umsetzung einer Mixed-Mode-Befragung mit postalischer Kontaktierung	115

Abbildung 19: Entwicklungsschritte einer Patientenbefragung und entsprechende Stellen
mit Anpassungsbedarf im Fall der Etablierung von Mixed-Mode-Befragungen 118

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AAPOR	American Association for Public Opinion Research
ALLBUS	Allgemeine Bevölkerungsumfrage der Sozialwissenschaften
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
CCA	corrected covered area
DeQS-RL	Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung
eGK	elektronische Gesundheitskarte
ESS	European Social Survey
Fragebogen-ID	Fragebogen-Identifikationsnummer
G-BA	Gemeinsamer Bundesausschuss
GVWG	Gesundheitsversorgungsweiterentwicklungsgesetz
IKT	Informations- und Kommunikationstechnologien
iOS	iPhone Operating System
ISCED 97	International Standard Classification of Education 1997
IVRS	Interactive Voice Response System
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
KI	Konfidenzintervall
LE	Leistungserbringer
NET	Nierenersatztherapie
NRSI	non-randomized studies of interventions
PDA	Personal Digital Assistant
PICO	Population, Intervention, Comparison, Outcome
PREM	Patient-Reported Experience Measure
PROM	Patient-Reported Outcome Measure
QI	Qualitätsindikator
QS	Qualitätssicherung
QS-Verfahren	Qualitätssicherungsverfahren

Abkürzung	Bedeutung
QS NET	Qualitätssicherungsverfahren <i>Nierenersatztherapie bei chronischem Nierenversagen einschließlich Pankreastransplantationen</i>
RCT	randomisierte kontrollierte Studie (<i>randomized controlled trial</i>)
SGB V	Fünftes Buch Sozialgesetzbuch
SHARE	Survey of Health, Ageing and Retirement in Europe
VPB	Versendestelle Patientenbefragung
VuMA Touchpoints	Verbrauchs- und Medienanalyse Touchpoints

Kurzfassung

Hintergrund

Patientenbefragungen spielen bei der Entwicklung von Qualitätssicherungsverfahren eine immer größere Rolle und bilden neben der QS-Dokumentation durch die Leistungserbringer und der Verwendung von Sozialdaten der Krankenkassen den dritten Baustein zur Abbildung der Versorgungsqualität. Die Entwicklungen von Patientenbefragungen des IQTIG basieren auf Patient-Reported Experience Measures (PREMs) und Patient-Reported Outcome Measures (PROMs), die Patientinnen und Patientinnen bislang als papierbasierte Fragebögen postalisch zu sich nach Hause geschickt bekommen („Paper-Pencil-Befragungen“). Vor dem Hintergrund der allgemeinen Forderung nach einem Ausbau von digitalen Anwendungen im Gesundheitswesen stellt sich die Frage, ob als Ergänzung oder Ablösung von papiergestützten Patientenbefragungen onlinebasierte Patientenbefragungen entwickelt und eingesetzt werden sollten. Hierbei gilt es Vorteile und Nachteile der verschiedenen Methoden abzuwägen. Generell könne onlinebasierte Befragungsmodi im Vergleich zu papierbasierten Befragungen mit Vorteilen einhergehen, z. B. durch die Möglichkeit zur unmittelbaren und damit aufwandsärmeren Erfassung der erhobenen Daten. Darüber hinaus bieten Onlinebefragungen methodische Vorteile, beispielsweise durch die Möglichkeiten zur Anwendung komplexer Filterführungen, ohne dabei die Fehleranfälligkeit bei der Beantwortung der Fragebögen zu erhöhen. Andererseits ergeben sich bei Onlinebefragungen spezifische methodische Herausforderungen, was insbesondere Fragen nach der Erreichbarkeit der Zielpopulation (Coverage-Problematik) und nach potenziellen Verzerrungen der Ergebnisse aufgrund von Moduseffekten in Bezug auf das Antwortverhalten betrifft. Weiterhin stellen sich im konkreten Kontext der Qualitätssicherung nach § 136 SGB V Fragen nach den gesetzlichen und datenschutzrechtlichen Rahmenbedingungen und den damit verbundenen grundsätzlichen Umsetzungsmöglichkeiten einer onlinebasierten Patientenbefragung als Datengrundlage zur Abbildung der medizinischen Versorgungsqualität.

Auftrag

Das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) wurde im Rahmen seiner Aufgaben nach § 137a Abs. 3 SGB V durch den G-BA am 20. November 2020 mit der „Entwicklung eines Konzeptes zur onlinebasierten Patientenbefragung“ beauftragt. Die Ziele der Beauftragung gliedern sich dabei in drei wesentliche Punkte: Erstens soll dargestellt werden, wie Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung zukünftig als Onlinebefragungen durchgeführt werden können, wobei unterschiedliche Befragungsmodi (Paper-Pencil-Befragungen, Onlinebefragungen, Mixed-Mode-Befragungen) hinsichtlich methodischer und umsetzungsrelevanter Aspekte unter Berücksichtigung der geltenden rechtlichen Rahmenbedingungen gegenübergestellt werden sollen. Zweitens soll erarbeitet werden, wie bereits bestehende oder in der Entwicklung befindliche Patientenbefragungen in eine Onlinebefragung überführt werden können. Drittens soll analysiert werden, welche Effekte und Konsequenzen mit der Umsetzung der Patientenbefragung als Onlinebefragung verbunden sind.

Zur Beantwortung der mit diesen Zielen einhergehenden Fragestellungen sollen gemäß Beauftragung Literaturrecherchen durchgeführt werden und verschiedene Befragungsmodi, ggf. mithilfe einer vergleichenden empirischen Untersuchung, gegenübergestellt werden. Zudem soll eine Prüfung der IT-technischen Umsetzungsmöglichkeiten unter Berücksichtigung der geltenden gesetzlichen und datenschutzrechtlichen Rahmenbedingungen erfolgen.

Als Ergebnis der Auftragsbearbeitung soll dargestellt werden, wie onlinebasierte Patientenbefragungen für den gegebenen Rahmen der sektorenübergreifenden Qualitätssicherung am besten umgesetzt werden können und welche Vor- und Nachteile im Vergleich zum bisherigen Konzept der Paper-Pencil-Befragung bestehen.

Methodisches Vorgehen

Auftragsgemäß wurden zur Beantwortung der mit den Projektzielen einhergehenden Fragestellungen Literaturrecherchen herangezogen. Da davon auszugehen ist, dass die auf Basis der Literatur identifizierten Effekte des Erhebungsmodus teilweise abhängig von den spezifischen Rahmenbedingungen und Inhalten einer Befragung sind – und daher nicht eindeutig auf den Kontext der externen Qualitätssicherung übertragbar sind – wurde zudem eine eigene empirische Untersuchung auf Basis von Primärdaten durchgeführt. Die einzelnen methodischen Blöcke werden im Folgenden näher dargestellt.

Zur Untersuchung der grundsätzlichen Erreichbarkeit potenzieller Zielgruppen von onlinebasierten Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung und zur Identifikation von potenziellen Selektionseffekten, die bei der Durchführung von Onlinebefragungen zu erwarten sind, wurde initial eine Recherche zu veröffentlichten Daten zur Internetnutzung in Deutschland durchgeführt. In die Analysen wurden neben Daten der amtlichen Statistik auch Ergebnisse von Erhebungen aus der Sozial- und Marktforschung einbezogen.

Zur Identifikation potenzieller Umsetzungsmöglichkeiten von onlinebasierten Befragungen, einschließlich der mit den unterschiedlichen Optionen einhergehenden Vor- und Nachteile sowie umsetzungsrelevanter Aspekte, wurde eine orientierende Literaturrecherche durchgeführt. Insbesondere berücksichtigt wurden dabei aktuelle Standardwerke der empirischen Sozialforschung mit dem Fokus auf Survey Methodology und Onlinebefragungen sowie Publikationen, die seit 2010 in thematisch relevanten wissenschaftlichen Fachzeitschriften veröffentlicht wurden. Die auf dieser Basis identifizierten Möglichkeiten zur Umsetzung von onlinebasierten Patientenbefragungen wurden hinsichtlich der IT-technischen bzw. prozessualen Umsetzbarkeit im Rahmen der gesetzlichen Qualitätssicherung sowie hinsichtlich der Vereinbarkeit mit geltenden gesetzlichen Rahmenbedingungen geprüft.

Um spezifisch für Befragungen zu gesundheitsbezogenen Themen zu analysieren, welche Effekte und methodische Konsequenzen bei der Umsetzung einer Onlinebefragung im Vergleich zu einer Paper-Pencil-Befragung zu erwarten sind, wurde eine systematische Literaturrecherche durchgeführt. Im Rahmen dieser Recherche wurden insbesondere die Fragen untersucht, inwiefern bei der Umstellung von einer papierbasierten auf eine onlinebasierte Patientenbefragung mit veränderten Rücklaufquoten zu rechnen ist und ob zu erwarten ist, dass das Antwortverhalten

von Patientinnen und Patienten zwischen unterschiedlichen Befragungsmodi variiert. Da auf Basis der vorab durchgeführten orientierenden Recherche davon ausgegangen werden konnte, dass bereits mehrere Übersichtsarbeiten zu diesen Themen publiziert wurden, wurde als Ansatz der systematischen Recherche eine Übersicht von systematischen Reviews („Overview of Systematic Reviews“) gewählt. Es wurden nur solche Arbeiten eingeschlossen, die kontrollierte Studien untersucht hatten und ab 2011 publiziert wurden.

Im Rahmen einer eigenen empirischen Untersuchung wurde zusätzlich der Einfluss unterschiedlicher onlinebasierter Befragungsmöglichkeiten auf Response-Raten (primäre Forschungsfrage), Teilnahmebereitschaft in Abhängigkeit von patientenseitigen Faktoren sowie Antwortverhalten und Datenqualität (sekundäre Forschungsfragen) im Vergleich zu den für die gesetzliche Qualitätssicherung herkömmlich entwickelten Papierfragebögen untersucht. Dabei wurden drei unterschiedliche onlinebasierte Umsetzungsmöglichkeiten der herkömmlichen papierbasierten Patientenbefragung in einem experimentellen Studiendesign gegenübergestellt: (1) „Single Mode Web“ (reiner Online-Fragebogen), (2) „Mixed-Mode“ (die Befragten haben die Möglichkeit, den Fragebogen entweder im Paper-Pencil-Format oder online auszufüllen), und (3) „Push-to-Web“ (die Befragten haben zuerst ausschließlich die Möglichkeit, den Fragebogen online auszufüllen; erst beim zweiten Erinnerungsschreiben wird die Möglichkeit geboten, alternativ einen Papierfragebogen zu beantworten). Nach der Festlegung der notwendigen Stichprobenfallzahlen zur Beantwortung der primären Forschungsfrage wurden Leistungserbringer (Dialyseeinrichtungen) zur Beteiligung an der Studiendurchführung rekrutiert.

Ergebnisse und Empfehlungen

Internetnutzung in Deutschland

Die Analysen der Daten zur Internetnutzung in Deutschland lassen den Schluss zu, dass mit einer reinen Onlinebefragung bzw. auf Basis einer elektronischen Kontaktaufnahme bestimmte Teile der Zielpopulation systematisch seltener erreicht werden könnten, als dies im Rahmen der herkömmlichen postalischen Paper-Pencil-Befragungen der Fall ist. Insbesondere Personen über 60 Jahren nutzen das Internet deutlich seltener als jüngere Personen. Darüber hinaus war ein niedriger Anteil an Internetnutzerinnen und Internetnutzern mit niedrigem Bildungsstand, weiblichem Geschlecht, dem Vorliegen von chronischen Erkrankungen sowie mit funktionellen Einschränkungen im Alltag assoziiert. Die Daten zur Internetnutzung in Deutschland zeigten zudem, dass es unter den Internetnutzerinnen und Internetnutzern in allen Altersgruppen einen nicht zu vernachlässigenden Anteil an Personen gibt, die E-Mail-Dienste nur selten oder gar nicht nutzen und dass dieser Anteil mit zunehmendem Alter und niedrigem Bildungsniveau steigt. Ähnliches gilt für die Nutzung von Smartphones. Dies verdeutlicht, dass hinsichtlich der Repräsentativität von onlinebasierten Befragungen mit Coverage-Problemen zu rechnen ist.

Identifikation und Bewertung der Umsetzungsmöglichkeiten

In der orientierenden Recherche zeigte sich hinsichtlich der Umsetzungsmöglichkeiten von onlinebasierten Patientenbefragungen, dass grundsätzlich zwischen Single-Mode-Onlinebefragungen und Mixed-Mode-Befragungen unterschieden werden kann. Eine Onlinebefragung im Single-Mode würde dabei die papierbasierte Patientenbefragung, wie sie aktuell im Rahmen der

gesetzlichen QS vorgesehen ist, gänzlich ersetzen. Bei Mixed-Mode-Befragungen bliebe die papierbasierte Befragung bestehen und würde um die Möglichkeit der Beantwortung des jeweiligen Fragebogens im Onlineformat ergänzt werden, je nach individueller Präferenz bzw. Vorliegen notwendiger Voraussetzungen seitens der Befragten. Mixed-Mode-Befragungen können wiederum in simultane und sequenzielle („Push-to-Web“) Ansätze unterteilt werden. Bei simultanen Mixed-Mode-Befragungen haben die Befragten bereits nach Erstkontakt die Möglichkeit, den Fragebogen entweder online oder postalisch als Papierfragebogen zu beantworten. Push-to-Web Befragungen hingegen zeichnen sich dadurch aus, dass die Befragten zuerst nur die Möglichkeit zur Onlinebefragung erhalten. Erst wenn die befragte Person nach einem gewissen Zeitraum den Fragebogen nicht online beantwortet hat, bekommt sie zusätzlich einen Papierfragebogen zugeschickt. Sowohl Single-Mode-Onlinebefragungen als auch Mixed-Mode-Befragungen scheinen auf Basis der Bewertung der Umsetzbarkeit und der gesetzlichen und datenschutzrechtlichen Rahmenbedingungen im Rahmen der gesetzlichen QS grundsätzlich umsetzbar. Beide Szenarien erfordern in diesem Zusammenhang geringfügige Anpassungen bei den Prozessen der Versendestelle der Patientenbefragung, die voraussichtlich relativ einfach zu implementieren sind. Im Gegensatz dazu ist aufgrund der Umstellung auf onlinebasierte Patientenbefragungen allerdings mit erheblichen Änderungen hinsichtlich der Aufgaben der Fragebogenannahmestelle zu rechnen, weshalb ggf. eine separate Fragebogenannahmestelle für onlinebasierte Patientenbefragungen eingerichtet bzw. das Aufgabenspektrum entsprechend erweitert werden müsste.

Ergebnisse der Übersicht von systematischen Reviews

In der Übersicht von systematischen Reviews konnten 16 Publikationen, die entsprechende Informationen zur Beantwortung der Forschungsfragen lieferten, eingeschlossen werden. Hinsichtlich der Frage nach unterschiedlichen Response-Raten je nach Erhebungsmodus zeigte sich ein heterogenes Bild. Während die Ergebnisse jener Studien, die Angaben zum Alter machen, darauf hindeuteten, dass insbesondere junge, nicht klinische Bevölkerungsgruppen ggf. besser mit Onlinebefragungen im Single-Mode zu erreichen sein können, fanden sich Hinweise dafür, dass in anderen Populationen elektronische bzw. onlinebasierte Befragungen mit niedrigeren Rücklaufquoten assoziiert sind, verglichen mit papierbasierten Befragungen. Weitere Studien wiederum fanden divergierende bzw. keine nennenswerten Unterschiede der Response-Raten zwischen den untersuchten Erhebungsmodi. Deutlich homogener hingegen fielen die Ergebnisse hinsichtlich der Übereinstimmung im Antwortverhalten je nach Befragungsmodus aus. Die überwiegende Mehrheit der eingeschlossenen Studien deutet demnach auf vergleichbare Ergebnisse zwischen onlinebasierten und papierbasierten Befragungen hin. Einschränkend anzumerken ist, dass nach kritischer Bewertung der überwiegende Teil der eingeschlossenen Reviews eine geringe methodische Qualität auf Basis der vorab definierten Anforderungen aufwies.

Ergebnisse der empirischen Untersuchung

Im Rahmen der empirischen Untersuchung konnte gezeigt werden, dass die Rücklaufquoten für die Einladungsform „Single Mode Paper“ am höchsten und für die Einladungsform „Single Mode Web“ am geringsten waren. Die Rücklaufquote der Einladungsform „Mixed-Mode“ unterschied sich nicht signifikant von der Rücklaufquote der Einladungsform „Single Mode Paper“. Bezüglich

der Datenqualität konnte eine höhere Datenqualität bei der Onlinebefragung festgestellt werden. So war die Wahrscheinlichkeit eine Frage nicht zu beantworten online geringer als in Papierform. Bezüglich der Vergleichbarkeit der Befragungsergebnisse wurden keine Unterschiede bezüglich des Modus gefunden, was mit den Ergebnissen der Übersicht von systematischen Reviews übereinstimmt.

Abgeleitete Empfehlungen

Auf Basis der Ergebnisse wurden die folgenden Empfehlungen für die Umsetzung von onlinebasierten Patientenbefragungen im Rahmen der gesetzlichen QS abgeleitet:

- Sowohl Single-Mode-Onlinebefragungen als auch Mixed-Mode-Befragungen scheinen im Rahmen der gesetzlichen QS grundsätzlich umsetzbar. Insbesondere aus Gründen der Zielerreichung und der erwarteten Rücklaufquoten empfiehlt das IQTIG jedoch einen Mixed-Mode Ansatz als präferierte Möglichkeit zur Durchführung von onlinebasierten Patientenbefragungen, bei der die Befragten selbst entscheiden können, ob sie die Befragung papierbasiert oder als Onlinefragebogen beantworten.
- Darüber hinaus wird auf Basis der gewonnenen Erkenntnisse empfohlen, für Patientenbefragungen im Rahmen der externen QS in aller Regel einen simultanen Mixed-Mode-Ansatz zu verfolgen, d. h., den Papier- und den Onlinefragebogen gleichzeitig zur Verfügung zu stellen.
- Für die Entwicklung von Patientenbefragungen im Rahmen von QS-Verfahren mit bestimmten Zielpopulationen (z. B. mit jüngeren, wenig komorbiden Patientinnen und Patienten) wird empfohlen, im Entwicklungsprozess zu prüfen, ob der Einsatz eines sequenziellen Mixed-Mode-Ansatzes („Push-to-Web“) oder aber eine Kombination aus beiden Mixed-Mode-Ansätzen zielführend ist.
- Im Fall einer simultanen Mixed-Mode-Befragung sollten in Anlehnung an die bisherigen Empfehlungen zur papierbasierten Patientenbefragung zum Zeitpunkt des Erstanschreibens neben den Zugangsdaten zur optionalen Onlinebefragung weiterhin Papierfragebögen, Anschreiben, vorfrankiertes Rücksendekuvert und Informationsschreiben postalisch zugesendet werden. Der Papierfragebogen sowie das zugehörige Rücksendekuvert sollten im Zuge des zweiten Erinnerungsschreibens erneut versendet werden. Aufgrund des sehr geringen Zusatzaufwands beim Versand sollten die Zugangsdaten für die Patientinnen und Patienten zur Onlinebefragung hingegen sowohl im Erstanschreiben als auch im Zuge beider Erinnerungsschreiben enthalten sein.
- Falls eine sequenzielle Mixed-Mode-Befragung („Push-to-Web“) durchgeführt werden soll, sollten die Patientinnen und Patienten zum Zeitpunkt des Erstanschreibens neben den Zugangsdaten zur Onlinebefragung nur ein Anschreiben und ein Informationsschreiben erhalten. Im Rahmen der beiden Erinnerungen sollten neben den Erinnerungsschreiben die Zugangsdaten zur Onlinebefragung erneut versendet werden. Zum Zeitpunkt des zweiten Erinnerungsschreibens sollten die Patientinnen und Patienten zusätzlich einen optionalen Papierfragebogen sowie ein zugehöriges vorfrankiertes Rücksendekuvert erhalten.
- Darüber hinaus wird empfohlen, dass die Kontaktaufnahme zu den Befragten grundsätzlich postalisch erfolgt. Dies gilt sowohl für die Erstanschreiben als auch für die beiden Erinnerungsschreiben.

- Für die Entwicklung einer Mixed-Mode-Befragung ist neben der Konzeption und Erstellung der Onlinebefragung und einer Zielgruppenanalyse mit Hinblick auf die Möglichkeit zur Push-to-Web-Befragung zusätzlich eine ausführliche Pretestung notwendig. Die Überführung bereits bestehender oder in der Entwicklung befindlicher Patientenbefragungen in eine Onlinebefragung erfordert daher eine eigene Weiterentwicklung.

Fazit und Ausblick

Gemäß Beauftragung durch den G-BA wurde auf Basis von Literaturrecherchen und eigenen empirischen Untersuchungen ein Konzept dafür entwickelt, wie onlinebasierte Patientenbefragungen im Rahmen der sektorenübergreifenden Qualitätssicherung zukünftig umgesetzt werden können und welche Konsequenzen damit assoziiert sind. Während verschiedene onlinebasierte Befragungsansätze als grundsätzlich umsetzbar bewertet wurden, spricht sich das IQTIG aus methodischen Gründen für Mixed-Mode-Befragungen mit postalischem Anschreiben als präferierten Ansatz von onlinebasierten Patientenbefragungen im vorgegebenen Rahmen aus. Die Vorteile gegenüber einer Onlinebefragung im Single-Mode begründen sich vorrangig damit, dass bei einer Mixed-Mode-Befragung insgesamt mit deutlich höheren Rücklaufquoten zu rechnen ist und unterschiedliche Bevölkerungsgruppen (z. B. hinsichtlich Alter, Bildungsstand und Krankheitsschwere) gleichermaßen erreicht werden können.

Die Vorteile einer Mixed-Mode-Befragung mit postalischem Anschreiben gegenüber der herkömmlichen papierbasierten Patientenbefragung im Single-Mode sind vielfältig. Da im Rahmen von Mixed-Mode-Befragungen ein Teil der Patientinnen und Patienten den Fragebogen online beantwortet und somit eine direkte Übermittlung von elektronischen Daten in Echtzeit gewährleistet werden kann, ist mit deutlichen Einsparungen von Ressourcen im Rahmen der Datenaufbereitung zu rechnen, aber auch mit niedrigeren Druck- und Portokosten, da weniger Erinnerungsschreiben verschickt werden müssen. Im Falle der Umsetzung einer Push-to-Web-Befragung würde es darüber hinaus bereits beim ersten Anschreiben zu einer deutlichen Reduktion der Druckkosten kommen. Zudem ist die Beantwortung der Fragebögen im Onlineformat aus vielfältigen Gründen mit einer besseren Datenqualität assoziiert, ohne dass dadurch mit einer systematischen Verzerrung der Befragungsergebnisse im Vergleich zu den Antworten auf Papierfragebögen zu rechnen ist.

Während die Umstellung auf Mixed-Mode-Befragungen mit postalischer Kontaktierung voraussichtlich nur mit geringfügigen Anpassungen bei den Prozessen der Versendestelle der Patientenbefragung einhergeht, ist mit erheblichen Änderungen hinsichtlich der Aufgaben der Fragebogenannahmestelle zu rechnen. Bevor Onlinebefragungen im Rahmen der gesetzlichen Qualitätssicherung zukünftig umgesetzt werden können, müsste aus Perspektive des IQTIG daher ggf. eine separate Fragebogenannahmestelle für onlinebasierte Patientenbefragungen eingerichtet bzw. das bisherige Aufgabenspektrum der Fragebogenannahmestelle deutlich erweitert werden. Darüber hinaus ist für die Überführung bereits bestehender oder in der Entwicklung befindlicher papierbasierter Patientenbefragungen in eine Onlinebefragung eine eigene Weiterentwicklung erforderlich.

Teil I: Hintergrund und Methoden

1 Einleitung

Die Perspektive der Patientinnen und Patienten bekommt eine zunehmende Relevanz bei der Abbildung der medizinischen Versorgungsqualität. Dies spiegelt sich insbesondere auch darin wieder, dass sich die rechtlichen Rahmenbedingungen in Deutschland in den letzten Jahren verändert haben und nun Patientenbefragungen als neue Datenquelle für die gesetzliche Qualitätssicherung zur Verfügung stehen. Damit spielen Patientenbefragungen bei der Entwicklung von Qualitätssicherungsverfahren (QS-Verfahren) eine immer größere Rolle, indem sie neben der Dokumentation durch die Leistungserbringer und der Verwendung von Sozialdaten der Krankenkassen den dritten Baustein zur Abbildung der Versorgungsqualität bilden.

In diesem Kontext hat der Gemeinsame Bundesausschuss (G-BA) im Jahr 2016 das IQTIG erstmals mit der Entwicklung von Patientenbefragungen für die QS-Verfahren *Perkutane Koronarintervention (PCI) und Koronarangiographie* und *Schizophrenie* („Versorgung von volljährigen Patientinnen und Patienten mit Schizophrenie, schizotypen und wahnhaften Störungen“) beauftragt (IQTIG 2018b, IQTIG 2018a). Im Jahr 2018 folgten die Beauftragungen zur Entwicklung von Patientenbefragungen für das QS-Verfahren *Nierenersatztherapie (NET: „Nierenersatztherapie bei chronischem Nierenversagen einschließlich Pankreastransplantationen“)* (IQTIG 2020) und für das QS-Verfahren zur ambulanten psychotherapeutischen Versorgung gesetzlich Krankenversicherter sowie 2019 zur Entwicklung eines sektorenübergreifenden, datengestützten Qualitätssicherungsverfahrens für Entlassmanagement, die auch eine Patientenbefragung enthält.

Die aktuellen Entwicklungen von Patientenbefragungen des IQTIG basieren auf Patient-Reported Experience Measures (PREMs) und Patient-Reported Outcome Measures (PROMs), die Patientinnen und Patientinnen als papierbasierte Fragebögen zu sich nach Hause geschickt bekommen („Paper-Pencil-Befragungen“). Neben den erforderlichen gesetzlichen Regelungen (z. B. zur Weitergabe der Patientenadressen), die im Rahmen des Krankenhausstrukturgesetzes (KHSKG) getroffen wurden, werden mit der Einrichtung einer Versendestelle für Patientenbefragungen und einer Fragebogenannahmestelle weitere strukturelle Maßnahmen zur Umsetzung der Fragebogenlogistik ergriffen.

Im Rahmen der Entwicklungen von Patientenbefragungen als Instrumente der gesetzlichen Qualitätssicherung stellt sich immer wieder die Frage, inwieweit der logistische Aufwand reduziert werden kann bzw. die Befragungen effizienter gestaltet und dabei möglichst viele Patientinnen und Patienten erreicht werden können. Insbesondere, was die Reduktion des Aufwands und damit die Schonung von Ressourcen anbelangt, scheinen onlinebasierte Befragungsmodi im Vergleich zu papierbasierten Befragungen mit einigen Vorteilen verbunden zu sein. Als einige Beispiele seien die unmittelbare Erfassung und Speicherung der erhobenen Daten, Möglichkeiten zur komplexen Filterführung und geringere Erhebungskosten zu nennen. Andererseits ergeben sich bei Onlinebefragungen spezifische methodische Herausforderungen und Fragen, wie die der Erreichbarkeit der Zielpopulation (Coverage-Problematik), der Teilnahmebereitschaft bzw.

Selbstselektion und folglich der Verzerrung der Ergebnisse, die je nach Zielpopulation und Befragungskontext variieren können. Weiterhin stellen sich im konkreten Kontext der Qualitätssicherung nach § 136 SGB V Fragen nach den gesetzlichen und datenschutzrechtlichen Rahmenbedingungen und den damit verbundenen grundsätzlichen Umsetzungsmöglichkeiten einer onlinebasierten Patientenbefragung als Datengrundlage zur Abbildung der medizinischen Versorgungsqualität.

Um abschätzen zu können, inwiefern sich onlinebasierte Patientenbefragungen für den gegebenen Rahmen der sektorenübergreifenden Qualitätssicherung eignen und welche Vor- und Nachteile im Vergleich zum bisherigen Konzept der Paper-Pencil-Befragung bestehen, müssen diese und weitere Fragestellungen geklärt werden.

2 Auftrag und Zielsetzung

Vor dem in Kapitel 1 beschriebenen Hintergrund wurde das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) im Rahmen seiner Aufgaben nach § 137a Abs. 3 SGB V durch den G-BA am 20. November 2020 mit der Entwicklung eines Konzepts zur onlinebasierten Patientenbefragung beauftragt (G-BA 2020). Als Ergebnis der Auftragsbearbeitung soll anhand einer abschließenden Empfehlung dargestellt werden, wie Onlinebefragungen im Rahmen der sektorenübergreifenden Qualitätssicherung zukünftig am besten umgesetzt werden können. Als übergeordnetes Ziel soll daher geprüft werden, inwiefern unterschiedliche Möglichkeiten einer onlinebasierten Umsetzung dazu beitragen können, Patientenbefragungen effizienter zu gestalten. Als Grundlage dafür soll eruiert werden, welche methodischen und ablauftechnischen Konsequenzen, Vorteile und Risiken mit onlinebasierten Patientenbefragungen einhergehen, mit welchen Auswirkungen hinsichtlich der Erreichbarkeit und Teilnahmebereitschaft der befragten Zielpopulation zu rechnen ist und welche rechtlichen und technischen Rahmenbedingungen bestehen.

Im Detail ergeben sich aus der Beauftragung des G-BA die folgenden Zielsetzungen:

- I. Darstellung, wie onlinebasierte Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung zukünftig durchgeführt werden können.
 - a. Identifikation und vergleichende Darstellung der Vor- und Nachteile der zur Verfügung stehenden Umsetzungsmöglichkeiten von Onlinebefragungen (z. B. Befragungsmodus im Mixed-Mode-Design oder als reine Onlinebefragung; Kontaktaufnahme per postalischem Anschreiben mit Link und / oder QR-Code, E-Mail oder SMS), einschließlich Möglichkeiten hinsichtlich Layout und Design sowie der Optionen von Erinnerungsverfahren (Reminder).
 - b. Darstellung der zu erwartenden prozessualen Veränderungen in der Umsetzung im Regelbetrieb durch die Umstellung auf Onlinebefragungen (z. B. Veränderung des Datenflusses, Veränderung organisatorischer Abläufe, spezifische Aufgaben der Versendestelle).
 - c. Darstellung der IT-technischen Anforderungen, die bei der Durchführung von Onlinebefragungen im Rahmen der gesetzlichen Qualitätssicherung zu berücksichtigen sind.
 - d. Darstellung der gesetzlichen und datenschutzrechtlichen Rahmenbedingungen bei der Umsetzung von Onlinebefragungen im Kontext der gesetzlichen Qualitätssicherung.
 - e. Darstellung der Veränderungen im Entwicklungsprozess und bei der Validierung von Patientenbefragungen bei der Umsetzung als Onlinebefragung.
- II. Empfehlungen dahingehend, wie bereits bestehende oder in Entwicklung befindliche Patientenbefragung in eine Onlinebefragung überführt werden können.

- a. Identifikation von Anforderungen an Layout und Design von Onlinebefragungen, die auf bestehenden Paper-Pencil-Fragebögen basieren.
 - b. Darstellung der notwendigen ergänzenden Entwicklungsschritte bei der Erstellung einer Onlineversion von bereits entwickelten Paper-Pencil-Fragebögen sowie einer Validierungsstrategie nachträglich entwickelter Onlinebefragungen.
- III. Analyse von Effekten und methodischen Konsequenzen, die mit der Umsetzung der Patientenbefragung als Onlinebefragung verbunden sind.
- a. Analyse der zu erwartenden Zielgruppenerreichung (Coverage) und der Selbstselektion bzw. Teilnahmebereitschaft bei der Umsetzung unterschiedlicher onlinebasierter Befragungsmodi im Vergleich zur herkömmlichen Paper-Pencil-Befragung, ggf. nach Subgruppen (z. B. Alter, Geschlecht, Gesundheitszustand, Bildung).
 - b. Analyse der Unterschiede im Antwortverhalten bei verschiedenen Befragungsmodi und Identifikation von zu erwartender Antwort- und Ergebnisverzerrungen durch den Umstieg auf eine Onlinebefragung.
 - c. Darstellung, ob und ggf. wie die Ergebnisse der onlinebasierten Patientenbefragung adjustiert werden sollen, um Vergleichbarkeit mit Paper-Pencil-Befragungen zu gewährleisten.

Zur Beantwortung der mit den Projektzielen einhergehenden Fragestellungen sollen laut Beauftragung Literaturrecherchen und – soweit notwendig – eine eigene empirische Untersuchung auf Basis von Primärdaten herangezogen werden. Die empirische Untersuchung soll gezielt und ressourcenschonend durchgeführt werden und auf bereits beauftragten Entwicklungen zur Patientenbefragung des G-BA aufbauen.

3 Orientierende Recherchen

3.1 Daten zur Internetnutzung in Deutschland

Zur Untersuchung der grundsätzlichen Erreichbarkeit der Zielgruppen von onlinebasierten Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung und zur Identifikation von potenziellen Selektionseffekten, die bei der Durchführung von Onlinebefragungen zu erwarten sind, wurde eine Recherche zu veröffentlichten Daten zur Internetnutzung in Deutschland durchgeführt. Zur Beantwortung der Fragestellung, ob sich bestimmte Subgruppen hinsichtlich der Erreichbarkeit über das Internet systematisch unterscheiden, erfolgten spezifische deskriptive Analysen der recherchierten Daten nach relevanten soziodemografischen Merkmalen. Entsprechend der Ergebnisse der orientierenden Recherche sowie auf Basis der Beauftragung des G-BA zählen hierzu vor allem Alter, Bildung, Geschlecht und Gesundheitszustand bzw. der funktionelle Status der Befragten.

In die Analysen wurden neben Daten der amtlichen Statistik auch Ergebnisse repräsentativer Erhebungen hinsichtlich (Subgruppen) der deutschen Allgemeinbevölkerung aus der Sozial- und Marktforschung einbezogen. Dies geschah vor dem Hintergrund, dass die publizierten Daten des Statistischen Bundesamtes zur Internetnutzung in Deutschland erstens keine hinreichende Differenzierung hinsichtlich der o. g. soziodemografischen Merkmale erlauben und zweitens nicht auf einer echten Zufallsstichprobe der deutschen Allgemeinbevölkerung basieren. Die Auswahl der recherchierten Datenquellen, die zur Analyse herangezogen wurden, erfolgte nach den Kriterien Aktualität, Art der Stichprobenziehung, Stichprobenumfang und der Verfügbarkeit interessierender Merkmale innerhalb der Datensätze. Nur bei echten Zufallsstichproben kann eine unverzerrte Schätzung der Internetnutzung sichergestellt werden. Damit die Unsicherheit nicht zu groß wird und die Ergebnisse sinnvoll interpretiert werden können, sollten die Stichprobenumfänge zudem nicht zu klein ausfallen. Die folgenden fünf Studien wurden schließlich in die Analysen aufgenommen:

Nutzung von Informations- und Kommunikationstechnologien (IKT)

Die IKT-Befragung stellt die Grundlage für die offiziellen Zahlen des Statistischen Bundesamtes zur Internetnutzung in Deutschland dar. Die Auswahl der teilnehmenden Haushalte erfolgt durch eine Quotenstichprobe mit freiwilliger Auskunftserteilung, die Befragung wird schriftlich durchgeführt. In der aktuellsten Erhebung für das Jahr 2020 wurden insgesamt 10.376 Haushalte befragt (Destatis 2021a).

Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS)

ALLBUS ist eine seit dem Jahr 1980 im Abstand von jeweils zwei Jahren durchgeführte sozialwissenschaftliche Befragung, deren inhaltlicher Schwerpunkt die Untersuchung von Einstellungen und Verhaltensweisen darstellt. Die aktuellsten Daten liegen für das Erhebungsjahr 2018 vor, für die Analysen wurden die Datensätze der Jahre 2012 (GESIS 2013), 2014 (GESIS 2018), 2016 (GESIS 2017) und 2018 (GESIS 2019) betrachtet, welche sämtlich auf Zufallsauswahlen der deut-

schen Allgemeinbevölkerung basieren. Die Ziehung der Stichproben erfolgt dabei aus den Registern der Einwohnermeldeämter. Die Befragung findet als persönlich-mündliches Interview statt, im Jahr 2018 konnten Interviews mit 3.477 Personen durchgeführt werden (GESIS 2019).

European Social Survey (ESS)

Der European Social Survey ist eine ländervergleichende sozialwissenschaftliche Befragung, welche seit 2002 ebenfalls im Abstand von zwei Jahren in vielen europäischen Ländern mit hohem methodischen Aufwand durchgeführt wird. Die aktuellsten verfügbaren Daten entstammen der Welle 9 und wurden in Deutschland im Zeitraum 2018/2019 erhoben. Auch die Stichprobe des ESS basiert auf einer Zufallsauswahl aus den Einwohnermeldeämtern, im Datensatz der Welle 9 sind Daten für 2.358 Personen enthalten (ESS Round 9 2021).

Survey of Health, Ageing and Retirement in Europe (SHARE)

ALLBUS und ESS haben den Nachteil, dass die relativ geringen Stichprobenumfänge eine weitere Differenzierung von Nutzungsgewohnheiten innerhalb der älteren Bevölkerung erschweren. Zudem fehlen in den sozialwissenschaftlichen Befragungen wichtige Informationen, insbesondere zum Gesundheitszustand der Befragten. Mit SHARE steht eine Längsschnittstudie zur Verfügung, welche sich mit den Themen Gesundheit und Lebensbedingungen der Bevölkerung in Europa ab einem Alter von 50 Jahren befasst (Börsch-Supan et al. 2013, Malter und Börsch-Supan 2017, SHARE-ERIC 2021). Zwar liegen die aktuellsten Daten zur Internetnutzung für das Jahr 2015 (Welle 6) (Börsch-Supan 2019) vor, die Studie verfügt jedoch für die hier zu beantwortenden Fragestellungen über einige wünschenswerte Eigenschaften. Hierzu gehört zunächst der Umstand, dass eine große Zufallsstichprobe von 4.412 Personen über 50 Jahren in Deutschland vorliegt und dass Personen in Alten- und Pflegeheimen nicht von der Befragung ausgeschlossen werden. Die Auswahlgrundlage bilden auch hier die Register der Einwohnermeldeämter. Zudem sind im Datensatz wertvolle Informationen zum Gesundheitszustand der Befragten enthalten, beispielsweise die Anzahl der ärztlich diagnostizierten Erkrankungen oder gesundheitlich bedingte Einschränkungen im Alltag. Trotz der Limitation, dass die Zahlen zur Internetnutzung möglicherweise aus heutiger Sicht eine Unterschätzung darstellen, scheinen die Daten der SHARE-Studie geeignet, um Unterschiede innerhalb der älteren Bevölkerung beispielsweise hinsichtlich des Gesundheitszustandes abbilden zu können.

VuMA Touchpoints

Während Daten zur Internetnutzung in allen erwähnten Studien erhoben werden, sind zuverlässige und aktuelle Informationen zur Verbreitung von Smartphones und der Häufigkeit der Verwendung von E-Mail-Diensten in Deutschland weniger leicht verfügbar. Aus diesem Grund wurde zusätzlich auf die Verbrauchs- und Medienanalyse (VuMa) Touchpoints, die seit 1995 von der Arbeitsgemeinschaft Verbrauchs- und Medienanalyse herausgegeben wird, zurückgegriffen (VuMa 2021). VuMa Touchpoints basiert zwar nicht auf einer Einwohnermeldeamtsstichprobe, sondern auf einem Random-Route-Verfahren, enthält jedoch einen breiten Fragenkatalog zum Thema Internet- und Mediennutzung. Der Stichprobenumfang ist zudem mit über 23.000 Personen im Befragungszeitraum zwischen 2018 und 2020 deutlich größer als in den genannten sozialwissenschaftlichen Erhebungen, was auch eine Betrachtung kleinerer Subgruppen ohne einen erheblichen Verlust an statistischer Präzision erlaubt.

Die Ergebnisse der Recherche nach Daten zur Internetnutzung in Deutschland sind in Kapitel 6 dargestellt.

3.2 Recherchen zur Umsetzungsmöglichkeiten von Onlinebefragungen

Zur Identifikation grundsätzlicher Umsetzungsmöglichkeiten von onlinebasierten Befragungen, einschließlich Aspekten hinsichtlich Layout und Design sowie der mit den unterschiedlichen Optionen einhergehenden Vor- und Nachteile und zur Ableitung entsprechender IT-technischer Anforderungen wurde eine orientierende Literaturrecherche durchgeführt. Insbesondere berücksichtigt wurden dabei aktuelle Standardwerke der empirischen Sozialforschung mit dem Fokus auf Survey Methodologie und Onlinebefragungen (Biffignandi und Bethlehem 2021, Biemer et al. 2017, Callegaro et al. 2015, Couper 2008, Dillman et al. 2014a, Groves et al. 2009, Schnell 2019, Schnell et al. 2018). Um die daraus abgeleiteten Themen inhaltlich zu ergänzen und mit aktuellen wissenschaftlichen Erkenntnissen abzugleichen wurden darüber hinaus die Titel und Abstracts aller veröffentlichten Publikationen gesichtet, die seit 2010 in den folgenden wissenschaftlichen Fachzeitschriften veröffentlicht wurden:

- Field Methods (<https://journals.sagepub.com/home/fmx>),
- International Journal of Public Opinion Research (<https://academic.oup.com/ijpor>),
- Journal of Official Statistics (<https://sciencdo.com/journal/jos>),
- Journal of Survey Statistics and Methodology (<https://academic.oup.com/jssam>),
- Journal of the American Statistical Association (<https://www.tandfonline.com/loi/uasa20>),
- Journal of the Royal Statistical Society (<https://rss.onlinelibrary.wiley.com/loi/14679876>),
- Kölner Zeitschrift für Soziologie und Sozialpsychologie (<https://kzfss.uni-koeln.de/de/>),
- Public Opinion Quarterly (<https://academic.oup.com/poq>),
- Social Science Computer Review (<https://journals.sagepub.com/home/ssc>),
- Sociological Methods & Research (<https://journals.sagepub.com/home/smr>),
- Survey Methodology (<https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X>),
- Survey Research Methods (<https://ojs.ub.uni-konstanz.de/srm>)

Die auf dieser Basis identifizierten Möglichkeiten zur Umsetzung von onlinebasierten Patientenbefragungen wurden hinsichtlich der IT-technischen bzw. prozessualen Umsetzbarkeit im Rahmen der gesetzlichen Qualitätssicherung sowie hinsichtlich der Vereinbarkeit mit geltenden gesetzlichen Rahmenbedingungen geprüft. Die Ergebnisse der orientierenden Recherche zu den Umsetzungsmöglichkeiten von onlinebasierten Befragungen sind in Kapitel 7 dargestellt.

4 Übersicht von systematischen Reviews

4.1 Zielsetzung der systematischen Recherche

Ein wichtiges Ziel des vorliegenden Projektes ist die Untersuchung von Effekten und den ggf. damit verbundenen methodischen Konsequenzen, die mit der Umsetzung der Patientenbefragung als Onlinebefragung einhergehen. Insbesondere soll analysiert werden, inwiefern bei der Umstellung von einer papierbasierten auf eine onlinebasierte Patientenbefragung mit veränderten Rücklaufquoten zu rechnen ist und ob zu erwarten ist, dass das Antwortverhalten von Patientinnen und Patienten bei selbstadministrierten Befragungen zu gesundheitsbezogenen Themen zwischen unterschiedlichen Befragungsmodi systematisch variiert.¹ Die systematische Literaturrecherche soll dementsprechend dazu dienen, sich gezielt den folgenden Fragen zu nähern:

- Welchen potenziellen Einfluss hat der Erhebungsmodus auf die Rücklaufquoten (Response-Raten) bzw. die Teilnahmebereitschaft bei selbstadministrierten Befragungen zu gesundheitsbezogenen Themen?
- Welchen potenziellen Einfluss hat der Erhebungsmodus auf das Antwortverhalten und die Befragungsergebnisse bei selbstadministrierten Befragungen zu gesundheitsbezogenen Themen?

Da auf Basis der vorab durchgeführten orientierenden Recherche (vgl. Kapitel 3) davon ausgegangen werden konnte, dass bereits mehrere Übersichtsarbeiten zu diesen Themen publiziert wurden, wurde im vorliegenden Projekt eine Übersicht von systematischen Reviews („Overview of Systematic Reviews“) gewählt. Dabei wurden explizit systematische Übersichtsarbeiten, die die Ergebnisse von Publikationen mit kontrollierten Studiendesigns zusammengefasst haben, eingeschlossen. Das methodische Vorgehen wird nachfolgend beschrieben. Die Ergebnisse der Übersicht von systematischen Reviews finden sich in Kapitel 8.

4.2 Präzisierung der Fragestellung

Zunächst erfolgte mithilfe des PICO-Schemas (Population, Intervention, Comparison, Outcome) die Operationalisierung in eine strukturierte, recherchierbare Fragestellung (Tabelle 1):

Welche Unterschiede hinsichtlich der Rücklaufquoten (Response-Raten) und des Antwortverhaltens (Messäquivalenz und Item-Nonresponse) zeigen sich zwischen verschiedenen selbstadministrierten Erhebungsmodi bei gesundheitsbezogenen Befragungen?

¹ Mit selbstadministrierten Befragungen sind solche Befragungen gemeint, die die Befragten ohne Anwesenheit einer Interviewerin bzw. eines Interviewers eigenständig beantworten können.

Tabelle 1: PICO-Schema

P	Klinische sowie nicht klinische Populationen
I	Jede Art der selbstadministrierten Erhebung von gesundheitsbezogenen Befragungsdaten
C	Eine andere Art der selbstadministrierten Erhebung von gesundheitsbezogenen Befragungsdaten
O	Rücklaufquote (Response-Rate), Übereinstimmung des Antwortverhaltens (Messäquivalenz und Item-Nonresponse)

4.3 Methodisches Vorgehen

Die Erstellung der Übersicht erfolgte in Anlehnung an die Kriterien für Overviews of Reviews aus dem Cochrane Handbook of Systematic Reviews of Interventions (Chapter V, Pollock et al. 2020). Ein entsprechendes Studienprotokoll wurde vorab im Open Science Framework registriert (Volland et al. 2020).

Das methodische Vorgehen wird nachfolgend beschrieben sowie etwaige Abweichungen zum a priori definierten Vorgehen dargelegt.

4.3.1 Einschlusskriterien

Eingeschlossen wurden systematische Reviews, die Primärstudien zum Vergleich zwischen verschiedenen selbstadministrierten Erhebungsmodi von gesundheitsbezogenen Befragungsdaten hinsichtlich zumindest einem der nachfolgenden Endpunkte untersuchten:

- Response-Rate
- Antwortverhalten

Um das Thema möglichst breit zu recherchieren, wurden als systematische Reviews alle Publikationen berücksichtigt, bei denen eine wissenschaftliche Synthese von Literatur erfolgte, die mittels Literaturrecherche und definierten Einschlusskriterien ausgewählt wurden. Eine kritische Bewertung der eingeschlossenen Studien war nicht zwingend gefordert. Allerdings erfolgte eine methodische Berücksichtigung dieses Aspekts bei der kritischen Bewertung der in dieser Übersicht eingeschlossenen systematischen Reviews (vgl. Abschnitt 4.6). Um einen möglichst aktuellen wissenschaftlichen Stand zu erfassen, erfolgte eine zeitliche Einschränkung der Suche auf die letzten 10 Jahre (d. h. ab 1. Januar 2011). Aus Effizienzgründen wurden zudem nur Volltexte in deutscher oder englischer Sprache berücksichtigt.

In Tabelle 2 sind die definierten Einschlusskriterien, die der systematischen Recherche und dem Screening der Artikel zugrunde lagen, aufgeführt.

Tabelle 2: Einschlusskriterien

	Einschluss
E1	Die Vollpublikation ist erhältlich.
E2	Die Publikationssprache ist Deutsch oder Englisch.
E3	Die Referenz wurde ab dem 1. Januar 2011 publiziert.
E4	Die Referenz ist keine Mehrfachpublikation.
E5	Die Publikation ist ein systematischer Review von kontrollierten Studien.
E6	Population: Klinische sowie nicht klinische Populationen.
E7	Intervention: Jede Art der selbstadministrierten Erhebung von gesundheitsbezogenen Befragungsdaten.
E8	Comparison: Eine andere Art der selbstadministrierten Erhebung von gesundheitsbezogenen Befragungsdaten.
E9	Outcome: Response-Rate, Übereinstimmung des Antwortverhaltens (Messäquivalenz und Item-Nonresponse)

4.4 Informationsquellen

4.4.1 Bibliographische Datenbanken

Am 1. Februar 2021 wurde eine systematische Literaturrecherche in den folgenden bibliographischen Datenbanken durchgeführt:

- Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) 1946 to January 29, 2021
- APA PsycInfo 1806 to January Week 4 2021 via Ovid
- Embase via Elsevier
- CINAHL via EBSCO
- Cochrane Database of Systematic Reviews (CDSR) via Wiley.

Ausgehend vom PICO-Schema und eines vorab vorliegenden Testsets aus relevanten, in MEDLINE indexierten Publikationen (n = 5) wurden die geeigneten Suchbegriffe abgeleitet. Für die Recherche wurde zunächst iterativ eine Strategie für die Literaturdatenbank MEDLINE entwickelt und dann entsprechend an die anderen Datenbanken angepasst. Die Suchstrategie bestand aus drei Blöcken:

1. Ein Rechercheblock, der die unterschiedlichen Arten der Datenerfassung bzw. Erhebungsmodi beschreibt (z. B. Online- oder Papierfragebogen).

2. Ein Rechercheblock, der die Datenerfassung auf selbstadministrierte Befragungsinstrumente zu gesundheitsbezogenen Themen beschränkt.
3. Einen Rechercheblock, der die Treffer auf systematische Reviews einschränkt.

Der verwendete Filter für die Eingrenzung der Treffer auf systematische Reviews stammt vom Scottish Intercollegiate Guidelines Network (SIGN [kein Datum]) und wurde geringfügig modifiziert sowie für die Datenbanken APA PsycInfo, Embase und CINAHL adaptiert. Alle Suchstrategien wurden hinsichtlich der inhaltlichen Vollständigkeit (Suchbegriffe, kontrolliertes Vokabular) sowie der Plattform-spezifischen Syntax von einer zweiten Person geprüft.

Bei der systematischen Suche konnten vier Publikationen aus dem Testset identifiziert werden. Der fünfte Artikel, welcher ebenfalls zur Ableitung der Suchwörter diente, ist thematisch passend und wurde inhaltlich mit dieser Suchstrategie identifiziert, konnte aber aufgrund des Publikationsdatums aus dem Jahr 2009 nicht mit der zeitlichen Einschränkung erfasst werden.

Folgende Limitationen wurden, falls in der jeweiligen Datenbank möglich, bei der Suchstrategie berücksichtigt:

- Publikationen ab 2011
- nur Publikationen in englischer oder deutscher Sprache.

Die jeweiligen Suchstrategien je nach Datenbank finden sich in Tabelle A 1, Tabelle A 2, Tabelle A 3, Tabelle A 4 und Tabelle A 5 im Anhang.

4.4.2 Handsuche in nicht indextierten Journalen

Um publizierte Forschung möglichst umfangreich zu recherchieren, wurde für bekannte und thematisch relevante Journale vorab geprüft, ob diese in den zu durchsuchenden bibliographischen Datenbanken indextiert sind. Im Ergebnis war eine manuelle Sichtung der Inhaltsverzeichnisse für vier Journale erforderlich:

- Journal of Official Statistics (<https://sciendo.com/journal/jos>)
- Survey Statistics and Methodology (<https://academic.oup.com/jssam>),
- Survey Methodology (<https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X>),
- Survey Research Methods (<https://ojs.ub.uni-konstanz.de/srm>)

Die Sichtung aller Ausgaben der vier seit 2011 publizierten Fachzeitschriften erfolgte bis zum 5. März 2021 auf den Webseiten der Verlage.

4.4.3 Citation searching

Weiterhin wurde für alle eingeschlossenen Artikel ein sogenanntes „citation searching“ durchgeführt. Im Detail wurden für jeden Artikel die folgenden Schritte durchgeführt:

- Prüfung der Referenzliste auf weitere relevante Literatur („backward citation“-Suche).
- Durchführung einer „forward citation“-Suche in Google Scholar.
- Durchführung einer „similar article“-Suche in PubMed, bei der die ersten 20 Treffer gesichtet wurden (Waffenschmidt et al. 2013).

Die einzelnen Schritte des „citation searching“ wurden am 13. März 2021 durchgeführt.

4.5 Selektionsprozess

Das Literaturmanagement der Titel und Abstracts der in den bibliographischen Datenbanken recherchierten Publikationen erfolgte mit dem Literaturverwaltungsprogramm EndNote. Die Deduplizierung erfolgte zum einen automatisiert mittels verschiedener Feld-Kombinationen in EndNote und zum anderen durch eine manuelle Sichtung eventueller Dubletten innerhalb aller Treffer sortiert nach Titel bzw. Autor.

4.5.1 Titel- und Abstract-Screening

Die deduplizierten Treffer der Suche in bibliographischen Datenbanken wurden zweifach unabhängig in Covidence (<https://covidence.org>) gesichtet und hinsichtlich ihrer inhaltlichen Relevanz bezogen auf die a priori festgelegte Frage überprüft und ausgewählt. Die Sichtung der Inhaltsverzeichnisse der vier, nicht in den durchsuchten bibliographischen Datenbanken indextierten Fachzeitschriften erfolgte von zwei Personen unabhängig voneinander auf den Webseiten der Verlage. Alle Treffer der verschiedenen Methoden des „citation searching“ wurden im PDF-Format bzw. mittels einer EndNote-Datenbank dokumentiert und zweifach unabhängig gesichtet. Uneinheitliche Bewertungen im Titel- und Abstract-Screening wurden diskutiert und eine Einigung herbeigeführt.

4.5.2 Volltext-Screening

Die Volltexte der ausgewählten Publikationen wurden in Covidence bzw. in EndNote von zwei Personen unabhängig voneinander dahingehend überprüft, ob die a priori festgelegten Einschlusskriterien (siehe Tabelle 2) zutreffen. Uneinheitliche Bewertungen wurden diskutiert und eine Einigung herbeigeführt. Die Dokumentation des Volltext-Screenings und der Ausschlussgründe für alle ausgeschlossenen Publikationen erfolgte in Covidence bzw. EndNote.

4.6 Kritische Bewertung

Mithilfe des AMSTAR 2-Instruments (A MeaSurement Tool to Assess systematic Reviews, Shea et al. 2017) erfolgte eine kritische Bewertung der berichteten methodischen Qualität der eingeschlossenen systematischen Reviews. Bei diesem Instrument sind 16 Items mit „Yes“, „Partial Yes“, „No“, „No meta-analysis conducted“ oder „Includes only NRSI“ bzw. „Includes only RCT“ zu beantworten (vgl. Tabelle A 6 im Anhang). Das Tool ist nicht auf die Bildung eines Gesamtwertes ausgerichtet. Shea et al. (2017) bieten aber ein Verfahren an, indem bestimmte Items (*critical domains*) als besonders entscheidend für die Qualität eines systematischen Reviews definiert werden können. Weist ein systematischer Review in einer dieser Items eine methodisch kritische Schwachstelle (*critical flaw*²) auf, ist der Grad des allgemeinen Vertrauens in die Ergebnisse als gering (*low*) anzusehen und bei mehr als einem *critical flaw* als *critically low* (sehr gering). Verfügt ein systematischer Review hingegen über keine kritischen Schwachstellen, hat

² Unter einem *critical flaw* wird das Bewerten einer kritischen Domäne mit *no* verstanden.

aber bei anderen (unkritischen) Items Schwächen (*non-critical weakness*³), ist das Gesamtvertrauen in die Ergebnisse als moderat (*moderate*) zu beschreiben. Bei lediglich einer bzw. keiner unkritischen Schwäche ist das Gesamtvertrauen in die Ergebnisse hoch (*high*) (Shea et al. 2017).

Anzumerken ist, dass das AMSTAR 2-Instrument explizit zur Bewertung von Übersichtsarbeiten, die Studien zu Gesundheitsinterventionen beinhalten, konzipiert wurde. Aufgrund der davon abweichenden Thematik und der Heterogenität der Studienlage in Bezug auf die Fragestellung im vorliegenden Projekt waren einzelne methodische Erfordernisse, wie sie für klassische Interventionsstudien im klinischen Bereich definiert sind, nicht zu erwarten. Daher wurden die Kriterien für die Bewertung einzelner Items für diese Übersicht von systematischen Reviews projektspezifisch angepasst. Eine Übersicht hierzu findet sich in Tabelle A 7 im Anhang. Die Items 2, 4, 9, 11, 13, 14, 15 wurden als *critical domains* eingestuft.

Die Bewertung der eingeschlossenen systematischen Reviews mittels AMSTAR 2 wurde von zwei Personen unabhängig voneinander durchgeführt und bei uneinheitlicher Bewertung diskutiert und konsentiert.

4.7 Datenextraktion

Ein Entwurf des Datenextraktionsformulars wurde erstellt, diskutiert und anschließend von zwei Personen pilotiert. Hierfür wurden für drei der eingeschlossenen systematischen Reviews die interessierenden Daten zweifach unabhängig extrahiert, abgeglichen und etwaige Unstimmigkeiten diskutiert. Im Ergebnis erfolgte die finale Modifizierung des Extraktionsformulars. Für alle eingeschlossenen systematischen Reviews wurden folgende Informationen, soweit verfügbar, extrahiert:

- Erstautor/Erstautorin, Publikationsjahr, Titel
- Informationen zur systematischen Suche: Anzahl und Name der durchsuchten Datenbanken sowie weiterer Informationsquellen, Suchzeitraum
- Charakteristika der befragten Population: Diagnose(n) und Alter (Mittelwert mit Standardabweichung, alternativ: Altersspanne oder sonstige Angaben), eingesetzte Fragebögen
- Ergebnisse des systematischen Reviews: Anzahl und Design der eingeschlossenen Studien (insgesamt), ob eine methodische Bewertung der eingeschlossenen Studien erfolgte
- Je Endpunkt: Art der untersuchten Datenerhebung (Intervention und Vergleich), Definition, Ergebnis (meta-analytisch oder deskriptiv), Anzahl der berücksichtigten Studien und deren Studiendesign, Anzahl analysierter Fälle, evtl. Krankheitsbilder der analysierten Teilnehmerinnen und Teilnehmer
- Benennung und Ausführung eventueller Interessenkonflikte der Autorinnen und Autoren sowie Benennung und Ausführung der Finanzierung des systematischen Reviews
- Ggf. Kommentare

Die Extraktion der weiteren eingeschlossenen Artikel wurde von einer Person durchgeführt und von einer weiteren Person geprüft. Kritische Punkte wurden diskutiert und im Konsens extrahiert. Alle Informationen wurden in der englischen Originalsprache der Artikel extrahiert.

³ Als *non-critical weakness* wird die Bewertung eines unkritischen Items mit *no* bezeichnet.

4.8 Überlappung

Für die Betrachtung des Überlappungsgrades der in den systematischen Reviews eingeschlossenen Publikationen wurde eine Zitationsmatrix nach Pieper et al. (2014) angelegt. Die Übersicht wurde von einer Person erstellt und von einer weiteren Person auf ihre Richtigkeit hin geprüft. Es wurde der korrigierte Überlappungsgrad (*corrected covered area*, CCA) nach Pieper et al. (2014) mit der folgenden Formel berechnet:

$$\frac{\text{Summe aller Studien aus den systematischen Reviews} - \text{Anzahl der Studien [Zeilen]}}{(\text{Anzahl der Studien [Zeilen]} * \text{Anzahl der systematischen Reviews [Spalten]}) - \text{Anzahl der Studien [Zeilen]}}$$

Ein CCA-Wert über 15 weist laut Pieper et al. (2014) auf eine sehr hohe Überlappung hin, während ein CCA-Wert bis 5 auf eine leichte Überlappung hindeutet.

4.9 Datensynthese

Es erfolgte eine deskriptive Beschreibung der systematischen Reviews, die zur Beantwortung der Fragestellung (vgl. Abschnitt 4.2) beitrugen. Eine Unterteilung erfolgte entsprechend der gefundenen Erhebungsmodalitäten der selbstadministrierten gesundheitsbezogenen Befragungen.

Die Ergebnisse bzgl. der Response-Rate und der Übereinstimmung des Antwortverhaltens wurden in Evidenztabelle dargestellt und im Text deskriptiv zusammengefasst. Dabei wurden die Ergebnisse sowie die methodische Qualität der systematischen Reviews berücksichtigt. Eine etwaige Überlappung von Studien in den eingeschlossenen systematischen Reviews und deren mögliche Auswirkungen auf die Datensynthese in dieser Übersichtsarbeit wurde erörtert.

5 Empirische Untersuchung

5.1 Zielsetzung und allgemeines Vorgehen

Es ist davon auszugehen, dass die bereits im Rahmen der vorab durchgeführten orientierenden Recherche (vgl. Abschnitt 3.2) identifizierten möglichen Effekte des Erhebungsmodus zumindest teilweise mit den spezifischen Rahmenbedingungen einer Befragung assoziiert sind. So kann insbesondere der Einfluss des Erhebungsmodus auf die Rücklaufquoten je nach Zielpopulation bzw. Zielsetzung der Befragung variieren. Aus diesem Grund sind die Ergebnisse der Literaturrecherche ggf. nicht uneingeschränkt bzw. nicht eindeutig auf onlinebasierte Patientenbefragungen im Kontext der externen Qualitätssicherung übertragbar. Um Effekte von unterschiedlichen onlinebasierten Befragungsmöglichkeiten insbesondere hinsichtlich Response-Raten, aber auch mit Hinblick auf Antwortverhalten bzw. Datenqualität im Vergleich zu den herkömmlich für die gesetzliche Qualitätssicherung entwickelten Paper-Pencil-Fragebögen zu untersuchen, wurde daher eine empirische Studie durchgeführt. Im Einzelnen sollten folgende Fragestellungen untersucht werden:

Primäre Forschungsfrage der empirischen Untersuchung

- Unterscheiden sich die Response-Raten bei verschiedenen Umsetzungsmöglichkeiten einer onlinebasierten Patientenbefragung von der herkömmlichen papierbasierten Patientenbefragung?

Sekundäre Forschungsfragen der empirischen Untersuchung

- Mit welchen patientenseitigen Faktoren (Alter, Geschlecht, Bildung, Multimorbidität) ist die Präferenz für einen bestimmten Erhebungsmodus assoziiert, wenn Patientinnen und Patienten die freie Wahl zwischen der Beantwortung eines Fragebogens in einem onlinebasierten oder einem papierbasierten Format haben?
- Gibt es Unterschiede hinsichtlich der Anzahl an Missings und „Weiß nicht mehr“-Antworten, die auf den Erhebungsmodus (online versus papierbasiert) zurückzuführen sind?
- Gibt es Unterschiede in den Ergebnissen der Befragung, die auf den Erhebungsmodus (online versus papierbasiert) zurückzuführen sind?

Um diese Forschungsfragen im Rahmen einer empirischen Studie zu untersuchen, wurden Dialyseeinrichtungen zur Beteiligung an der Studiendurchführung rekrutiert. Es folgten die Ausschreibung für die Unterstützung bei der Datenerhebung an externe Dienstleister sowie die Rekrutierung der zu befragenden Patientinnen und Patienten. Anschließend fand die eigentliche Datenerhebung (Feldphase) statt, inklusive der Einladungen zur Teilnahme an der Studie sowie die administrative Begleitung der Studie. Nach dem Ende der Datenerhebung wurden die erhobenen Daten durch das IQTIG ausgewertet (siehe Kapitel 9).

Zur Untersuchung der o. g. Fragestellungen mussten zuerst Patientenfragebögen, die zuvor vom IQTIG als Paper-Pencil-Befragung zum Zweck der externen Qualitätssicherung entwickelt und validiert wurden, in eine Onlinebefragung übertragen werden. Um Moduseffekte möglichst zu vermeiden, wurde darauf geachtet, dass sich das Design der Onlinebefragung am Design der

Paper-Pencil-Fragebögen orientiert hat. Darüber hinaus wurden bei der Programmierung gängige Empfehlungen zum Design von Onlinebefragungen berücksichtigt (siehe Abschnitt 7.2). Die Onlinebefragung wurde so konzipiert, dass sie auf PC, Tabletcomputer und Smartphone unter Verwendung gängiger Webbrowser durchgeführt werden konnte und dabei den geltenden Datenschutzstandards entspricht.

Die Programmierung der Onlinebefragung auf Basis bereits entwickelter Paper-Pencil-Fragebögen, der Versand und die Annahme der Fragebögen bzw. der Links zur Onlinebefragung sowie die Dateneingabe wurden von einem vom IQTIG beauftragten externen Dienstleister übernommen. Vor Beginn der Rekrutierung der Patientinnen und Patienten wurde ein Ethikvotum der International Medical & Dental Ethics Commission GmbH (IMDEC) zur Durchführung des Forschungsvorhabens eingeholt. Nach Begutachtung des Studienprotokolls wurde der Durchführung der Studie zugestimmt.

5.2 Zielpopulation und verwendete Fragebögen

Laut Beauftragung sollten die durchgeführten Analysen gezielt und ressourcenschonend durchgeführt werden und dementsprechend auf den bereits beauftragten Entwicklungen zur Patientenbefragung des G-BA aufbauen (G-BA 2020). Dies betraf insbesondere die Nutzung von bereits entwickelten Fragebögen zur Durchführung der empirischen Studie. Bisher liegen validierte Patientenfragebögen für die QS-Verfahren *Perkutane Koronarintervention und Koronarangiographie (QS PCI)* und *Schizophrenie* („Versorgung von volljährigen Patientinnen und Patienten mit Schizophrenie, schizotypen und wahnhaften Störungen“) sowie für das QS-Verfahren *Nierenerersatztherapie (QS NET)* vor.

Für die geplante Studie im vorliegenden Projekt wurden die bereits entwickelten Fragebögen für Patientinnen und Patienten mit ständiger Dialyse aus der Patientenbefragung *QS NET* (Fragebogen A-e „Dialyse in einer Einrichtung“ und Fragebogen A-h „Heimdialyse“) (IQTIG 2020) aus verschiedenen Gründen für besonders geeignet befunden, weshalb diese Patientengruppe als Zielpopulation definiert wurde. Neben dem Vorliegen von validierten Fragebögen, die zum Zweck der gesetzlichen Qualitätssicherung entwickelt wurden, besteht der größte Vorteil dieser Patientengruppe darin, dass mit relativ großen Fallzahlen innerhalb eines vergleichsweise kurzen Rekrutierungszeitraumes zu rechnen ist. Dies hat den Hintergrund, dass die Beantwortung der genannten Fragebögen aus dem Verfahren *QS NET* nicht auf einen ganz bestimmten Moment im Rahmen der Behandlung festgelegt ist, sondern die Fragebögen zu nahezu jedem beliebigen Zeitpunkt im Rahmen der ständigen Dialysebehandlung eingesetzt werden können (IQTIG 2020). Darüber hinaus lassen die Daten des Standard-Pretests der Patientenbefragung *QS NET* darauf schließen, dass die Zusammensetzung einer Stichprobe von Patientinnen und Patienten mit Dialyse dazu geeignet ist, potenzielle patientenseitige Einflussvariablen auf die Beantwortung von Onlinebefragungen – wie Komorbiditäten oder Alter (vgl. Kapitel 6) – zu untersuchen. So findet sich beispielsweise unter dialysepflichtigen Patientinnen und Patienten nicht nur ein hoher Anteil an älteren Personen und Personen mit kardiovaskulären Erkrankun-

gen, sondern mit rund 13 % gaben im Standard-Pretest auch relativ viele Patientinnen und Patienten an, dass von einer Ärztin oder einem Arzt eine Depression festgestellt worden sei, sodass auch Patientinnen und Patienten mit psychischen Erkrankungen adressiert sind.

5.3 Studiendesign

Basierend auf den Ergebnissen der orientierenden Recherchen (vgl. Kapitel 3, 6 und 7) sollten drei unterschiedliche onlinebasierte Umsetzungsmöglichkeiten jeweils der herkömmlichen papierbasierten Patientenbefragung („Single Mode Paper“) in einem experimentellen Studiendesign gegenübergestellt werden, wobei die Items bzw. der Inhalt der Befragung in allen Umsetzungsmöglichkeiten identisch war. Damit ergaben sich vier Studienarme:

Single Mode Paper (Kontrollbedingung): Die Befragten bekamen postalisch einen Papierfragebogen zugeschickt, der anhand des beiliegenden Rücksendeumschlags postalisch retourniert werden kann.

Single Mode Web: Die Befragten wurden postalisch angeschrieben und erhielten im Anschreiben einen Link und QR-Code, um den Online-Fragebogen aufzurufen und online auszufüllen.

Mixed-Mode: Die Befragten wurden postalisch angeschrieben, erhielten einen Papierfragebogen (inkl. Rücksendekuvert) und zusätzlich wurde im Anschreiben ein Link und QR-Code angegeben, um den Fragebogen online auszufüllen. Somit hatten die Befragten die Möglichkeit, den Fragebogen entweder im Paper-Pencil-Format oder online auszufüllen.

Push-to-Web: Die Befragten wurden postalisch angeschrieben und erhielten im Anschreiben einen Link und QR-Code, um den Fragebogen aufzurufen und online auszufüllen. Hatten die Befragten nach vier Wochen den Fragebogen noch nicht ausgefüllt, erhielten sie mit dem zweiten und damit letzten Erinnerungsschreiben (siehe Abschnitt 5.7) zusätzlich zum Link zum Onlinefragebogen einen Papierfragebogen inkl. Rücksendeumschlag zugeschickt, um den Fragebogen dann je nach Präferenz als Papierversion beantworten zu können.

In allen vier Studienarmen erfolgte die Kontaktierung der Teilnehmerinnen und Teilnehmer postalisch, wobei die Teilnehmerinnen und Teilnehmer zum Zeitpunkt der Einwilligung zur Studienteilnahme nicht wussten, zu welchem der vier Studienarme sie schlussendlich per Zufallsauswahl zugeteilt werden würden. Ferner wurden die potenziellen Teilnehmerinnen und Teilnehmer vorab darüber aufgeklärt, dass sie per Zufallsauswahl den Fragebogen entweder als Papierversion oder im Onlineformat zugeschickt bekommen würden, oder aber selbst aus einer der beiden Möglichkeiten wählen können. Auf einen zusätzlichen Studienarm mit elektronischer Kontaktierung, beispielsweise über Smartphones oder E-Mail-Adressen, wurde aus mehreren Gründen verzichtet:

- Auf Basis der Ergebnisse der vorab durchgeführten orientierenden Recherche (vgl. insbesondere Kapitel 6) ist nicht davon auszugehen, dass mittelfristig auf Basis von elektronischen Kontaktierungsmöglichkeiten alle relevanten Subpopulationen hinsichtlich Alter, Geschlecht, Bildungsstatus und Gesundheitszustand der jeweiligen Zielpopulation der Patientenbefragung gleichermaßen erreicht werden können.

- Der Grundsatz der Datensparsamkeit erfordert, dass die Verarbeitung zusätzlicher Patientendaten – wie E-Mail-Adressen oder Telefonnummern – sehr gut begründet werden müsste. Dem gegenüber stehen die Ergebnisse der Literatur (vgl. Kapitel 7 und 8), dass postalische bzw. papierbasierte Befragungen und Mixed-Mode-Befragungen mit vielen Vorteilen, insbesondere hinsichtlich der Teilnahmebereitschaft und Response-Raten, einhergehen. Bei Mixed-Mode-Befragungen scheint eine generelle postalische Kontaktierung die einzige praktikable Umsetzungsmöglichkeit.
- Aus Praktikabilitätsgründen scheint es nahezu ausgeschlossen, dass die einzelnen Leistungserbringer im späteren Regelbetrieb die notwendigen Informationen (z. B. gültige E-Mail-Adressen, gültige Telefonnummern bei Smartphone-Nutzung) aller ihrer Patientinnen und Patienten einholen und diese zur weiteren Verarbeitung weiterleiten.
- Darüber hinaus ist es wichtig, dass die experimentellen Rahmenbedingungen zwischen den Studienarmen konstant gehalten werden, um valide Vergleiche zwischen den Studienarmen zu gewährleisten. Dies trifft insbesondere auf Faktoren zu, die die grundsätzliche Einwilligung bzw. Möglichkeit zur Studienteilnahme potenziell beeinflussen. Nur so können Effekte auf die jeweilige Umsetzungsmöglichkeit der onlinebasierten Befragung zurückgeführt werden. Bei einem Studienarm, der die zusätzliche Einholung einer elektronischen Kontaktmöglichkeit zur Teilnahme voraussetzt, wäre davon auszugehen, dass es zu einer starken Vorselektion hinsichtlich teilnahmebereiter Personen mit besonderer Affinität zu Onlinemedien käme, während Patientinnen und Patienten ohne Internetzugang bzw. ohne Smartphone oder E-Mail-Adresse gar nicht erst die Teilnahmebedingungen für diesen Studienarm erfüllen würden. Die ermittelten Response-Raten, die auf Basis der Teilnehmerinnen und Teilnehmer mit aktiver Einwilligung zur Studienteilnahme ermittelt werden, könnten somit nicht sinnvoll mit den anderen Studienarmen verglichen werden.

5.4 Allokation und Randomisierung

Bei einer klassischen randomisierten Kontrollstudie („randomized controlled trial“, RCT) erfolgt eine einfache randomisierte Zuteilung bzw. Allokation der Studienteilnehmerinnen und -teilnehmer zu den Studienarmen. Allerdings ist es häufig sinnvoll, eine geschichtete bzw. stratifizierte Zuteilung vorzunehmen, wenn zum einen die Fallzahlen eher klein sind und zum anderen der Verdacht besteht, dass ein relevantes Merkmal aufseiten der Patientinnen und Patienten mit dem Ergebnis der Testung zusammenhängt (Kernan et al. 1999). Da die Internetnutzung stark mit dem Alter zusammenhängt (vgl. Kapitel 6) wurde eine randomisierte Zuteilung zu den Armen proportional geschichtet nach Altersquartilen vorgenommen, um die Präzision zu erhöhen und den Einfluss von Störvariablen zu reduzieren. Durch die Stratifizierung des Alters nach Quartilen ergaben sich 4 Altersgruppen.

Zusätzlich zu den Altersgruppen wurde auch nach der Dialyseeinrichtung (bzw. dem Leistungserbringer, LE), in der die Teilnehmerinnen und Teilnehmer behandelt wurden, geschichtet. Dadurch sollte verhindert werden, dass potenzielle (nicht gemessene) einrichtungsspezifische Effekte die Response-Raten einzelner Studienarme bzw. den Effekt des Erhebungsmodus verzerren. Somit konnte sichergestellt werden, dass pro Einrichtung jede Altersgruppe in allen experimentellen Armen vertreten war. Ferner konnte somit auch bei den späteren Analysen mit

höherer Sicherheit festgestellt werden, ob ein Effekt des Erhebungsmodus tatsächlich auf diesen zurückzuführen ist und nicht auf Merkmale der einzelnen Leistungserbringer. Einen Überblick zum geplanten Design und zur randomisierten Zuteilung bietet Abbildung 1.

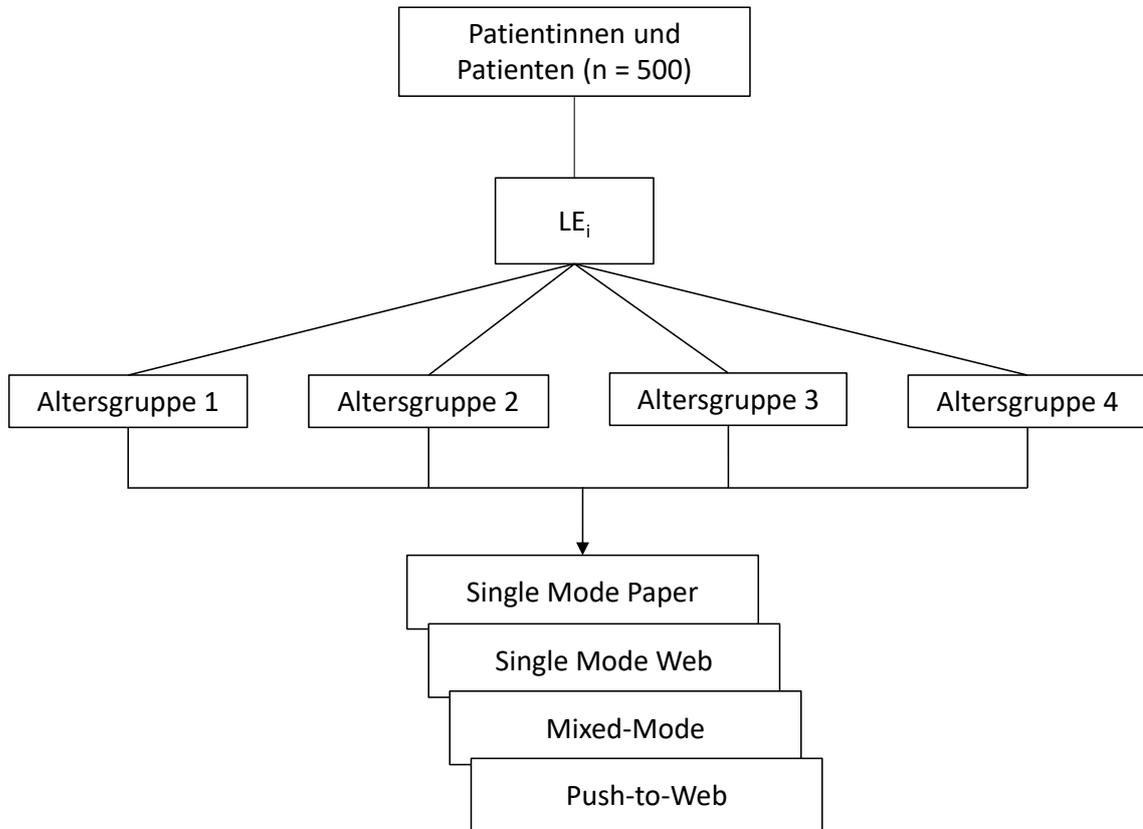


Abbildung 1: Design der empirischen Testung

5.5 Fallzahlkalkulation

Die Fallzahlkalkulationen wurden in Bezug auf die primäre Forschungsfrage nach den unterschiedlichen Rücklaufquoten je nach Studienarm durchgeführt. Für die Bestimmung der nötigen Stichprobengröße zur Feststellung von signifikanten Unterschieden der Response-Raten zwischen jeweils zwei Studienarmen wurden Poweranalysen für den zweiseitigen Test auf Unterschied eines Anteilswert in zwei Populationen in R gerechnet (R Core Team 2021). Für die Analysen mussten Annahmen hinsichtlich der zu erwartenden Response-Raten für den Kontrollarm (Single Mode Paper) angenommen werden. Dieser Wert wurde auf Basis der Erfahrungen mit postalischen Befragungen von Dialysepatientinnen und -patienten im Rahmen des Standard-Pretests der Patientenbefragung *QS NET* geschätzt und dementsprechend auf 85 % Fragebogenrücklauf festgelegt (IQTIG 2020). Ferner wurde für den Fehler 1. Art (Signifikanzniveau Alpha) ein Wert von 0,05 und für den Fehler 2. Art (Beta-Fehler) ein Wert von 0,8 festgesetzt.

Weiterhin wurde festgelegt, dass eine Differenz von mindestens 15 Prozentpunkten zwischen der Response-Rate des Armes „Single Mode Paper“ und den anderen experimentellen Armen detektiert werden soll (Blumenberg und Barros 2018). Entsprechend den auf dieser Basis durch-

geführten Poweranalysen werden pro Studienarm ca. $n = 125$ Patientinnen und Patienten benötigt, um mit der gewünschten Power feststellen zu können, ob einer der Versuchsarme bei der Response-Rate einen Unterschied von 15 Prozentpunkten zu der erwarteten Response-Rate des Kontrollarms von 85 % aufweist. Deswegen kalkuliert das IQTIG mit einer Gesamtfallzahl von $n = 500$. Einen Überblick über die Ergebnisse der Poweranalysen bieten Abbildung 2. Hier wird zudem deutlich, dass die benötigten Fallzahlen pro experimentellen Arm in Abhängigkeit von der Differenz der angenommenen Response-Raten der zu vergleichenden Studienarme stark sinken. Je größer der Unterschied in den Response-Raten, desto weniger Patientinnen und Patienten werden benötigt.

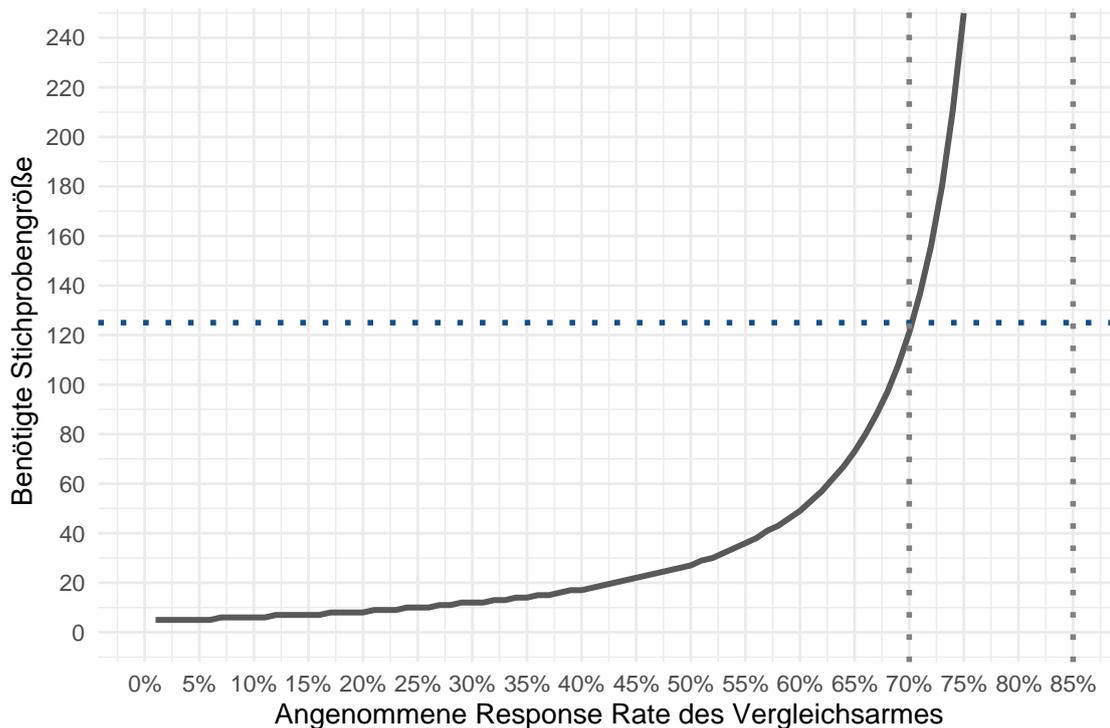


Abbildung 2: Angenommene Response-Rate des Vergleichsarmes und benötigte Stichprobengröße

Da die Poweranalysen im Hinblick auf die Beantwortung der primären Forschungsfrage berechnet wurden, kann es je nach Höhe des tatsächlichen Fragebogenrücklaufs vorkommen, dass die Power hinsichtlich der sekundären Forschungsfragen zu gering ist, um kleinere Effekte zu detektieren.

5.6 Rekrutierung der Teilnehmerinnen und Teilnehmer

Bezüglich der Rekrutierungskriterien wurden für die empirische Untersuchung volljährige und einwilligungsfähige Patientinnen und Patienten mit regelmäßiger Dialyse seit mindestens drei Monaten (ambulant oder teilstationär) eingeschlossen. Aufgrund der Tatsache, dass die verwendeten Fragebögen für gesetzlich versicherte Patientinnen und Patienten entwickelt wurden, sollten möglichst keine privatversicherten Personen zur Studienteilnahme rekrutiert werden. In Ausnahmefällen, wie z. B. bei Rekrutierungsschwierigkeiten in kleineren Dialyseeinrichtungen, durften nach Rücksprache mit dem IQTIG auch privat versicherte Patientinnen und Patienten

mitaufgenommen werden, da grundsätzlich keine Hinweise dafür vorliegen, dass der Versichererstatus einen unmittelbaren Einfluss auf die Bearbeitung des jeweiligen Fragebogens hat.

Die Rekrutierung der befragten Patientinnen und Patienten erfolgte durch teilnehmende Dialyseeinrichtungen. Diese wurden im Vorfeld durch einen Aufruf auf der Website des IQTIG sowie durch die Nutzung von fachspezifischen E-Mail-Verteilern dazu eingeladen, freiwillig an der Studie teilzunehmen. Im Aufruf wurde darum gebeten, für einen befristeten Zeitraum die Adressdaten von Patientinnen und Patienten, die entsprechend der Rekrutierungskriterien zur Studienteilnahme geeignet waren und sich freiwillig zur Teilnahme bereiterklärten, zur Verfügung zu stellen. Die Dialyseeinrichtungen wurden im Aufruf zur Studienteilnahme auch über den Hintergrund und den geplanten Ablauf der Studie, über entsprechende Teilnahmevoraussetzungen sowie über den Umgang mit Daten und die Gewährleistung des Datenschutzes aufgeklärt. Schlussendlich haben sich 16 Dialyseeinrichtungen für die Teilnahme an der Studie registriert, die eine große Bandbreite hinsichtlich Träger, Größe sowie Standort und Region (urban vs. rural) aufwiesen. Alle 16 Einrichtungen wurden eingeladen, an der Studie teilzunehmen.

Die Rekrutierungsphase startete am 15. März 2021. Das Vorgehen bei der Rekrutierung der Patientinnen und Patienten durch die Dialyseeinrichtungen wurde vom IQTIG vorgegeben. Insbesondere wurden die Dialyseeinrichtungen darauf hingewiesen, bei der Ansprache von Patientinnen und Patienten keine Vorauswahl nach eignen Kriterien vorzunehmen.⁴ Die Einrichtungen informierten zuerst ihre Patientinnen und Patienten über die Studie und fragten nach der Teilnahmebereitschaft. Gaben die Patientinnen und Patienten nach einer vollständigen Information über das Vorhaben und eine Aufklärung über die Datenverwendung und den Datenschutz ihr schriftliches Einverständnis zur Teilnahme an der Studie, erfassten die Leistungserbringer anhand einer ihnen durch das IQTIG zur Verfügung gestellten Microsoft-Excel-Vorlage Adressdaten und ergänzende Informationen.⁵ Diese ergänzenden Daten wurden u. a. für die Nonresponse-Analyse und für die Auswahl des spezifischen zu versendenden Fragebogens benötigt. Das IQTIG stellte den Leistungserbringern für die Patientinnen und Patienten zusätzlich Informationsmaterial in Form von Postern und Flyern zur Verfügung, in denen die Studie vorgestellt wurde und die Kontaktpersonen aus dem IQTIG benannt wurden. Die Leistungserbringer luden die Microsoft-Excel-Tabelle mit den Informationen der Patientinnen und Patienten spätestens zum zuvor festgelegten Datenlieferzeitpunkt (16. April 2021) auf eine durch einen externen Dienstleister bereitgestellte sichere Plattform hoch.

⁴ Für die Aussagekraft der Studie war es von hoher Bedeutung, dass alle Patientinnen und Patienten, die grundsätzlich in der Lage sind, einen Fragebogen auszufüllen, die gleiche Chance hatten, in die Studie miteinbezogen zu werden. Daher sollten explizit auch ältere Patientinnen und Patienten mit eingeschlossen werden, wie auch Patientinnen und Patienten, die keinen Internetanschluss bzw. keine Erfahrung mit / kein Interesse an Onlinemedien haben. Die rekrutierenden Einrichtungen wurden dafür sensibilisiert, dass nur so ein valider Vergleich der Antwortbereitschaft bei unterschiedlichen Befragungsmodi stattfinden kann.

⁵ Für Patientinnen und Patienten in Dialysebehandlung wurden neben der Anrede, ggf. Titel, Name, Adresse sowie Geburtsjahr und Geschlecht zusätzlich das Datum des Beginns der regelmäßigen Dialysebehandlung in der Einrichtung sowie die Dialysemodalität und -organisationsform von den Leistungserbringern erfasst.

5.7 Feldphase

Die erhobenen Daten der teilnehmenden Patientinnen und Patienten wurden von den Dialyseeinrichtungen über eine sichere Datenverbindung an den externen Dienstleister übermittelt. Dazu erhielt jede teilnehmende Einrichtung einen persönlichen Zugang zu einer Onlineplattform. Ein Exemplar der unterzeichneten Einwilligungserklärungen wurden von den Leistungserbringern postalisch an den externen Dienstleister gesendet, ein zweites Exemplar verblieb bei den Studienteilnehmerinnen und -teilnehmern.

Der Versand der Fragebögen wurde in enger Absprache mit dem IQTIG durch den externen Dienstleister durchgeführt. Dem Fragebogen lagen ein Anschreiben mit Datum der Einwilligung und der Nennung der rekrutierenden Dialyseeinrichtung sowie ein Informations- und Datenschutzblatt und ein vorfrankierter Rückumschlag bei. In dem Anschreiben wurde auch auf Ansprechpartnerinnen und Ansprechpartner und eine Telefonhotline des IQTIG hingewiesen.

Je nach Studienarm wurden die Papierfragebögen oder nur die Anschreiben mit den Links zur Onlinebefragung am 7. Mai 2021 versandt. Um die Rücklaufquote zu erhöhen, wurde an alle Teilnehmerinnen und Teilnehmer zwei Wochen nach Fragebogenversand ein kombiniertes Erinnerungs- und Dankeschreiben geschickt. Dort wurde den Personen gedankt, die den Fragebogen bereits ausgefüllt zurückgeschickt hatten und gleichzeitig wurden die Personen, die dies noch nicht getan hatten, an die Befragung erinnert. In Anlehnung an die Empfehlungen von Dillman et al. (2014b) wurde weitere zwei Wochen später ein zweites Erinnerungsschreiben an jene Teilnehmerinnen und Teilnehmer gesendet, bei denen der Eingang ihres Fragebogens mittels kontrollierten Rücklaufs bis zu diesem Zeitpunkt noch nicht verzeichnet werden konnte. In den Studienarmen „Single Mode Paper“ und „Push-to-Web“ wurde im Rahmen des zweiten Erinnerungsschreibens ein Papierfragebogen mitgeschickt.

Auf den Anschreiben wie auch auf den Erinnerungsschreiben war jeweils ein Rücksendedatum angegeben, bis wann der ausgefüllte Fragebogen spätestes zurückgeschickt werden sollte. Die späteste Fragebogenannahme erfolgte am 25. Juni 2021. Nach dem Ende der Feldphase wurden dem IQTIG die anonymisierten Datensätze seitens des externen Dienstleisters übermittelt.

5.8 Statistische Analysen

Zur Überprüfung der Unterschiede der Response-Rate in den jeweiligen experimentellen Armen im Vergleich zum Kontrollarm (primäre Forschungsfrage der empirischen Untersuchung) wurden Chi²-Tests und Cramer's V für kategoriale Daten berechnet. Um eine potenzielle Verzerrungen durch Unit-Nonresponse zu untersuchen, wurde eine binomial-logistische Regression gerechnet.

Zur Überprüfung, ob in den experimentellen Armen „Mixed-Mode“ und „Push-to-Web“ die Präferenz der Teilnehmerinnen und Teilnehmer für einen bestimmten Erhebungsmodus (online versus papierbasiert) mit Alter, Bildung, Geschlecht oder Multimorbidität assoziiert war, wurden t-Tests für metrische Variablen und exakte Tests nach Fisher für die kategorialen Variablen berechnet.

Bezüglich der Unterschiede der Antwortqualität zwischen der onlinebasierten und der papierbasierten Erhebungsform – gemessen an der Anzahl an Missings und „Weiß nicht mehr“-Antworten – wurden binomial-logistische Regressionsmodelle mit den Kontrollvariablen Alter, Geschlecht, Bildung sowie Anzahl der Komorbiditäten herangezogen. Die abhängigen Variablen waren jeweils der Anteil der Missings bzw. „Weiß nicht mehr“-Antworten und die unabhängige Variable der Erhebungsmodus (onlinebasiert versus papierbasiert).

Zur Untersuchung systematischer Unterschiede bei der Beantwortung der PREMs und PROMs aus der Patientenbefragung *QS NET*, die auf den Erhebungsmodus (onlinebasiert versus papierbasiert) zurückzuführen sind, wurden ebenfalls binomial-logistische Regressionsmodelle herangezogen. Die abhängigen Variablen stellten dabei Personen-Scores jener PREMs und PROMs dar, die zur Berechnung von 13 befragungsbasierten Qualitätsindikatoren (QI) empfohlen wurden.⁶ Für die Berechnung der Personen-Scores wurde aus den zugehörigen Fragebogen-Items eines bestimmten QIs zuerst ein Summenscore pro Patientin bzw. Patient gebildet. Dieser Summenscore wurde dann ins Verhältnis zum maximal möglichen Summenscore gesetzt.⁷ Somit konnten Personen-Scores auch im Fall von fehlenden Werten bei einzelnen Items berechnet werden. Die Bildung solcher Scores auf Patientenebene („Personen-Scores“) wurde als methodisch gerechtfertigt erachtet, da die Items der Qualitätsindikatoren hinsichtlich ihrer internen Konsistenz und Homogenität mindestens zufriedenstellende Werte aufwiesen (IQTIG 2020). Bei den Berechnungen wurden dieselben Kontrollvariablen wie bei den Regressionen zur Antwortqualität in die Modelle mit aufgenommen.

Bei Signifikanztests wurde ein p-Wert von $< 0,05$ als statistisch signifikant betrachtet. Cramer's V Werte ab 0,1 wurden als kleiner Effekt, ab 0,3 als mittlerer und ab 0,5 als starker Effekt interpretiert (Cohen 1988). Bei dem Vorliegen von Überdispersion wurden statt den binomial-logistischen Regressionsmodellen quasibinomial-logistische Regressionsmodelle berechnet.

⁶ Die 13 QIs beziehen sich jeweils auf Patientinnen und Patienten mit ständiger Dialyse. Ausgenommen wurden hierbei vier weitere befragungsbasierte QIs für den Bereich Dialyse, die sich ausschließlich auf Patientinnen und Patienten mit Erstdialyse beziehen, sowie ein weiterer QI, der sich ausschließlich auf Patientinnen und Patienten mit Heimdialyse bezieht. Hintergrund ist, dass die zu erwartenden Fallzahlen im Rahmen des vorliegenden Studiendesigns zu klein ausgefallen wären, um aussagekräftige Schlüsse hinsichtlich dieser Qualitätsindikatoren ziehen zu können.

⁷ Ein Personen-Score wird dabei so gebildet, dass die einzelnen Werte der Antwortoptionen der Fragebogen-Items aufaddiert und dann ins Verhältnis zum maximal möglichen Summenscore gesetzt werden. Für ein Beispiel mit zwei Items, deren Antwortoptionen jeweils die Ausprägungen „Ja“ = 1 und „Nein“ = 0 aufweisen, bedeutet dies, dass eine Person, die beide Items mit „Ja“ beantwortet hat, als Summe einen Wert von 2 erhält. Wird dieser Wert dann ins Verhältnis zum maximal möglichen Summenscore gesetzt, erhält die Person schließlich einen Personen-Score von $2 / 2$. Sollte die Person auf einem der beiden Items ein Missing haben und das andere mit „Ja“ beantwortet haben, wird der maximal mögliche Summenscore entsprechend auf 1 angepasst. Der Personen-Score wäre in diesem Fall $1 / 1$.

Teil II: Ergebnisse

6 Nutzung des Internets und von Onlinemedien in Deutschland

Zur Sicherstellung einer repräsentativen Abbildung der patientenberichteten Behandlungsqualität ist es erforderlich, dass alle zur Zielpopulation gehörenden Patientinnen und Patienten die gleiche Wahrscheinlichkeit aufweisen, an der Befragung teilnehmen zu können. Neben der grundsätzlichen Bereitschaft der Befragten an der Befragung teilzunehmen muss daher für eine erfolgreiche Durchführung einer Onlinebefragungen der Zugang zum Medium Internet bzw. die generelle Nutzung von Onlinemedien in allen zur Zielpopulation gehörenden Subgruppen gewährleistet sein. Werden einzelne Gruppen durch fehlenden Zugang ausgeschlossen bzw. systematisch seltener erreicht, so muss vor allem dann mit verzerrten Ergebnissen gerechnet werden, wenn das interessierende Merkmal mit der Wahrscheinlichkeit zur Teilnahme an der Befragung korreliert (Schnell 2019). In diesem Kapitel wird auf Basis verschiedener zur Verfügung stehender Datenquellen zur Internetnutzung in Deutschland untersucht, ob eine Nutzung des Internets in unterschiedlichen Bevölkerungsgruppen gleichermaßen vorausgesetzt werden kann und mit welchen Selektionseffekten im Kontext von Patientenbefragungen auf dieser Basis möglicherweise gerechnet werden muss.

6.1 Internetnutzung in Deutschland

Das Statistische Bundesamt veröffentlicht im Rahmen der Erhebung „Nutzung von Informations- und Kommunikationstechnologien“ (IKT) regelmäßig offizielle Daten zur Internetnutzung in Deutschland (Destatis 2021b). Die durch das Statistische Bundesamt für das Jahr 2020 ermittelten Daten zur Internetnutzung sind in Tabelle 3 dargestellt.⁸ Für Personen bis einschließlich 64 Jahre zeigt sich in den Daten eine Nutzung des Internets von nahezu 100 %, in der Altersgruppe der Über-64-Jährigen fällt der Anteil auf 69 % ab.

Tabelle 3: Internetnutzung in Deutschland, Quelle: Destatis (2020) (Destatis 2021b)

Alter	Anteil Internetnutzerinnen und Internetnutzer
16–24 Jahre	100 %
25–44 Jahre	100 %
45–64 Jahre	97 %
65+ Jahre	69 %

Für eine systematische Untersuchung des Zugangs zum Internet in unterschiedlichen soziodemografischen Gruppen sind die verfügbaren aggregierten Statistiken des Statistischen Bundesamtes nicht geeignet. Zum einen fasst das Statistische Bundesamt alle Personen über 65 Jahren in einer einzigen Kategorie zusammen, was eine weitere Differenzierung in dieser Altersgruppe

⁸ Die Einteilung der Altersklassen wurde an die Kategorisierung des Statistischen Bundesamtes angepasst.

nicht erlaubt. Zum anderen stellt die IKT selbst keine Zufallsstichprobe dar und basiert auf freiwilligen Auskünften. Da sich die Befragung ausschließlich mit dem Thema IKT beschäftigt, ist davon auszugehen, dass jene Personen, welche die adressierten Medien nutzen, in der Befragung überrepräsentiert sind. Dies wird auch im Methodenbericht der IKT limitierend angemerkt (Destatis 2021a).

Unabhängig davon, ob die in Tabelle 3 genannten Zahlen eine Überschätzung des tatsächlichen Anteils darstellen, lässt der geringere Anteil der Internetnutzerinnen und Internetnutzer in der Altersgruppe der Über-64-Jährigen deutlich erkennen, dass ältere Personen über das Internet schwerer erreicht werden können. Darüber hinaus lässt der starke Rückgang in der höchsten Altersgruppe die Frage offen, welche Faktoren (z. B. Bildung oder gesundheitlicher Status) in dieser Subpopulation mit der Nutzung des Internets assoziiert sind.

Eine detailliertere Betrachtung der Internetnutzung bzw. des Zugangs zu Onlinemedien in den höheren Altersgruppen ist für die Beantwortung der Fragestellungen im vorliegenden Projekt notwendig. Gemeinsam mit der Tatsache, dass die Daten der IKT aus den oben genannten Gründen vermutlich eine Überschätzung der tatsächlichen Internetnutzung in Deutschland darstellen, erscheint es daher sinnvoll, auf weitere Datenquellen zurückzugreifen, um ein umfassendes Bild der Internetnutzung in relevanten Subgruppen gewinnen zu können. Solche Datensätze sollten bestenfalls echte Zufallsstichproben der deutschen Allgemeinbevölkerung darstellen, einen ausreichend großen Stichprobenumfang aufweisen und über einen Fragenkatalog verfügen, welcher Aufschlüsse über die Internetnutzung innerhalb relevant erscheinender soziodemografischer Subgruppen zulässt. Hierzu zählen die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) (GESIS 2019), der European Social Survey (ESS) (ESS Round 9 2021), der Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan 2019) und auch die Verbrauchs- und Medienanalyse (VuMA) Touchpoints (VuMa 2021). Die jeweils aktuellsten verfügbaren Daten zur Internetnutzung liegen für die Jahre 2018 bzw. 2019 (ALLBUS/ESS/VuMA) sowie das Jahr 2015 (SHARE) vor.⁹ Die Auswahlgrundlage für ALLBUS, ESS und SHARE bilden Daten der Einwohnermeldeämter, die Stichprobe der VuMA Touchpoints basiert auf einem Random-Route-Verfahren.¹⁰ Bei allen Erhebungen handelt es sich um persönliche Befragungen in Anwesenheit einer das Interview führenden Person. Der Einbezug mehrerer Studien soll das Risiko falscher Schlussfolgerungen aufgrund der Betrachtung einer einzigen Stichprobe und Fragestellung minimieren.

Die Frage nach der Internetnutzung ist in den betrachteten Studien unterschiedlich operationalisiert. Eine Übersicht der jeweils gestellten Fragen, Antwortmöglichkeiten und der Codierung für die vorgenommenen Analysen ist in Tabelle 4 dargestellt.

⁹ Für SHARE sind Daten der Welle 7 (2017) zwar bereits verfügbar, Informationen zur Internetnutzung wurden jedoch zuletzt in Welle 6 (2015) für alle Befragten erhoben.

¹⁰ Bei einem Random-Route-Verfahren handelt es sich um eine Art von Zufallsauswahl. Will man z. B. eine Zufallsstichprobe der Einwohnerinnen und Einwohner Berlins, hat aber keine vollständige Liste der Haushalte, so würde man bei dem Random-Route-Verfahren eine zufällige Startadresse auswählen. Ausgehend davon würde eine Interviewerin bzw. ein Interviewer einer zufällig generierten Begehungsanweisung folgen, welche zu einem Wohnhaus führt. Leben in dem Wohnhaus mehrere Haushalte, wird per Zufallsauswahl ein Haushalt gezogen und dann interviewt. Mit diesen Verfahren lassen sich Zufallsstichproben beispielsweise von Haushalten ziehen, ohne eine vollständige Liste zu haben (Schnell et al. 2018).

Tabelle 4: Fragen zur Internetnutzung in den Studien ALLBUS, ESS, VuMA Touchpoints und SHARE

ALLBUS (Welle 2018)	
Frage	Das Internet gewinnt eine immer größere Bedeutung in der Gesellschaft. Deshalb interessiert uns, ob Sie selbst zumindest gelegentlich das Internet für private Zwecke nutzen? (Intervieweranweisung: Sei es mittels eines Computers, Laptops, Tablets oder Smartphones)
Antwortkategorien	1 Ja, ich nutze das Internet für private Zwecke 2 Nein, ich nutze das Internet nicht für private Zwecke
Codierung	1 = Internetnutzung
ESS (Welle 2018/19)	
Frage	Man kann das Internet auf verschiedenen Geräten nutzen, wie beispielsweise Computer, Tablets und Smartphones. Wie oft nutzen Sie das Internet auf so einem oder einem anderen Gerät, sei es bei der Arbeit oder für private Zwecke?
Antwortkategorien	1 Nie 2 Nur ab und zu 3 Ein paarmal pro Woche 4 An den meisten Tagen 5 Jeden Tag
Codierung	2–5 = Internetnutzung
VuMA Touchpoints (Welle 2018/19)	
Frage	Haben Sie in den letzten 12 Monaten das Internet genutzt, egal ob zu Hause, am Arbeitsplatz, in der Schule, Universität oder an einem anderen Ort? Denken Sie bitte auch daran, dass man inzwischen mit einer Vielzahl von Geräten das Internet nutzen kann. Also neben PCs, Laptops oder Netbooks, auch mit Tablets, Smartphones oder internetfähigen Fernsehern oder Spielekonsolen.
Antwortkategorien	1 Ja 2 Nein
Codierung	1 = Internetnutzung
SHARE (Welle 2015)	
Frage	Haben Sie während der letzten 7 Tage das Internet mindestens einmal genutzt, sei es für E-Mails, zur Informationssuche, für Einkäufe oder zu einem anderen Zweck?
Antwortkategorien	1 Ja 2 Nein
Codierung	1 = Internetnutzung

In Abbildung 3 ist der Anteil an Internetnutzerinnen und Internetnutzern entsprechend der in Tabelle 4 dargestellten Codierungen nach Altersgruppen vergleichend für den ALLBUS, den ESS und die VuMA Touchpoints dargestellt. Aufgrund des abweichenden Erhebungszeitraums und einer deutlich unterschiedlichen Fragestellung werden die Daten des SHARE zunächst nicht eingeschlossen, da eine unmittelbare Vergleichbarkeit der Daten nicht gegeben ist.

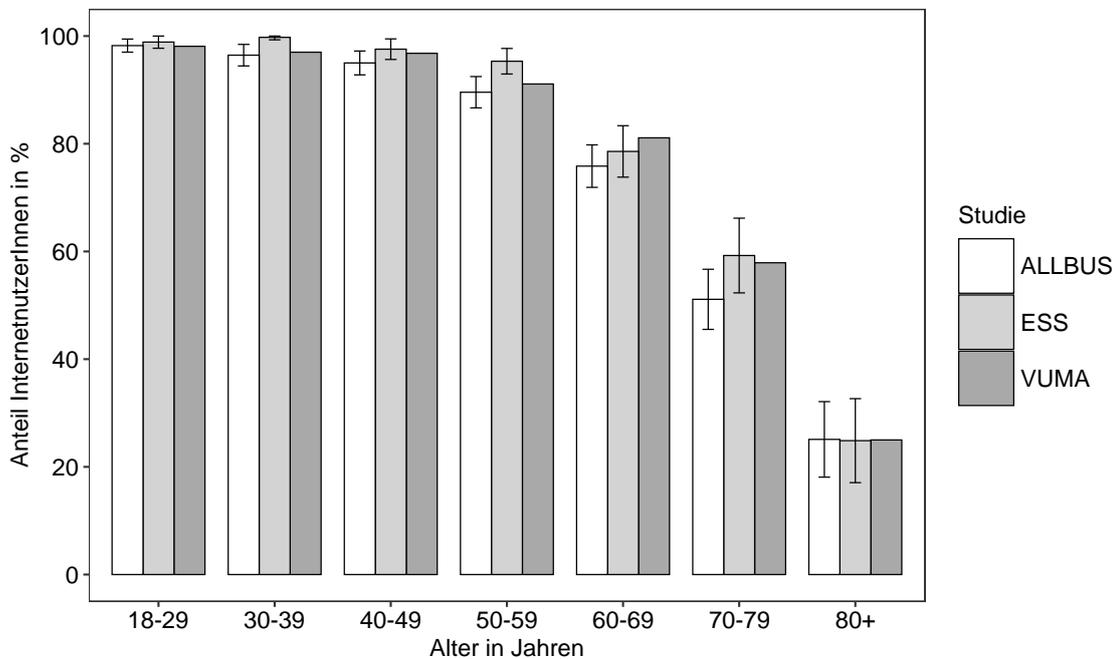


Abbildung 3: Anteil Internetnutzerinnen und -nutzer in Deutschland 2018/19, Datenbasis: ESS (ESS Round 9 2021) / ALLBUS (GESIS 2019) / VuMA Touchpoints (VuMa 2021)¹¹

Obwohl sich der Fragewortlaut und die Art der Stichprobenziehung zwischen den drei Studien unterscheiden, zeigt sich hinsichtlich der Internetnutzung in den dargestellten Altersgruppen über alle Studien hinweg ein bemerkenswert konsistentes Bild. Während eine nahezu flächendeckende Nutzung des Internets in den Altersgruppen bis zu 60 Jahren zu beobachten ist, zeigt sich ein starker Rückgang der Internetnutzung in den Altersgruppen über 60 Jahren. Besonders deutlich wird der Zusammenhang zwischen Alter und Internetnutzung bei den Über-80-Jährigen, von denen in den Erhebungsjahren 2018 und 2019 in etwa nur 25 % das Internet nutzten. Der Umstand, dass die Gruppe der Über-80-Jährigen laut amtlicher Statistik im Jahr 2020 etwa 5,68 Millionen Menschen umfasste (Eurostat 2021) macht deutlich, dass hinsichtlich der Repräsentativität einer onlinebasierten Befragung mit einem erheblichen Coverage-Problem zu rechnen ist. Dies gilt in besonderem Maße für Befragungen von Patientinnen und Patienten, da hier die Altersverteilung der Zielpopulation in vielen Versorgungsbereichen verglichen mit der Allgemeinbevölkerung deutlich höher ausfallen dürfte.

¹¹ Für die VuMA Touchpoints können keine Konfidenzintervalle angegeben werden, da die Daten lediglich gewichtet und in aggregierter Form zur Verfügung stehen.

Zwar hat die Nutzung des Internets im Zeitverlauf innerhalb der letzten Jahre in allen Altersgruppen kontinuierlich zugenommen, die Geschwindigkeit dieses Prozesses sollte jedoch nicht überschätzt werden. In Abbildung 4 ist die Entwicklung des Anteils der Internetnutzerinnen und Internetnutzer für die Altersgruppen ab 40 Jahren zwischen den Jahren 2012 und 2018 auf Basis der Daten des ALLBUS dargestellt.¹² In den Altersgruppen ab 60 Jahren zeigt sich im betrachteten Zeitraum von sechs Jahren ein Anstieg um etwa 15 bis 20 Prozentpunkte, die erhebliche Kluft zwischen jüngeren und älteren Personen bleibt jedoch deutlich erkennbar. Auch wenn davon auszugehen ist, dass sich der Trend zu einer häufigeren Nutzung des Internets auch in der älteren Bevölkerung fortsetzen wird, scheinen vernachlässigbare Unterschiede zwischen den Altersgruppen in den kommenden Jahren nicht wahrscheinlich. Neben dem allgemeingültigen Umstand, dass es für valide Ergebnisse in der Zielpopulation keine systematischen Unterschiede beim Zugang zum Erhebungsmedium geben sollte, kann sich bei einer geringen Erreichbarkeit älterer Personen über das Internet auf Leistungserbringerebene zusätzlich das Problem ergeben, dass schlicht eine zu geringe Fallzahl zustande kommt, um zuverlässige Aussagen über die Qualität der Behandlung zuzulassen. Dies ist besonders dann zu erwarten, wenn ein Großteil der Behandelten innerhalb einer Einrichtung ein höheres Alter aufweist.

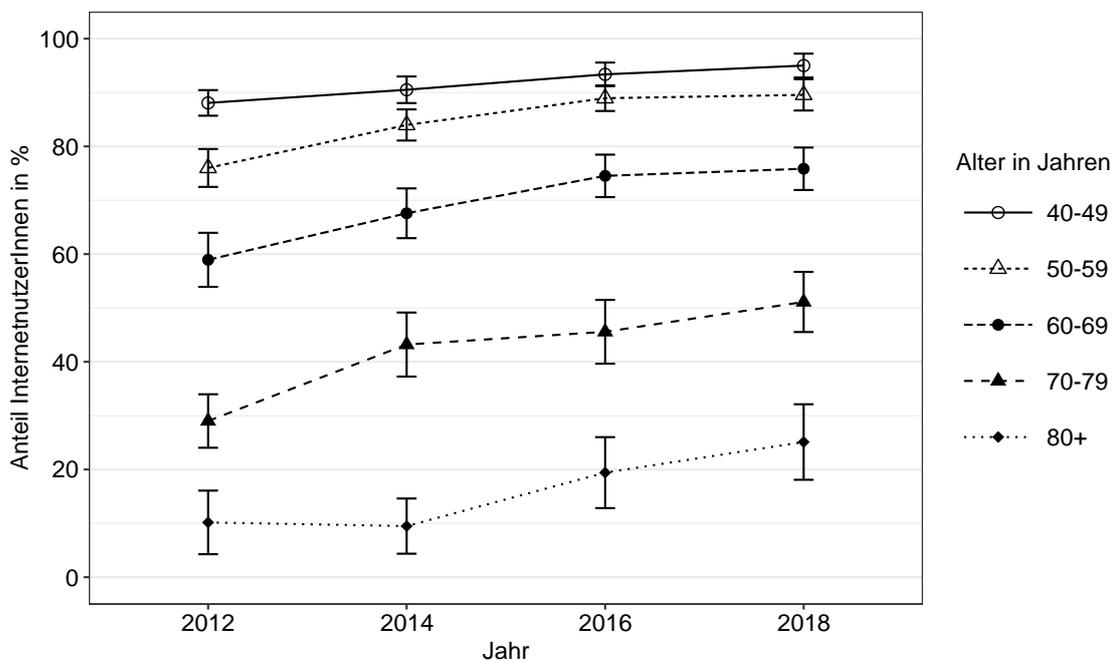


Abbildung 4: Internetnutzung in Deutschland im Zeitverlauf, Datenbasis: ALLBUS 2012 (GESIS 2013), 2014 (GESIS 2018), 2016 (GESIS 2017) und 2018 (GESIS 2019)

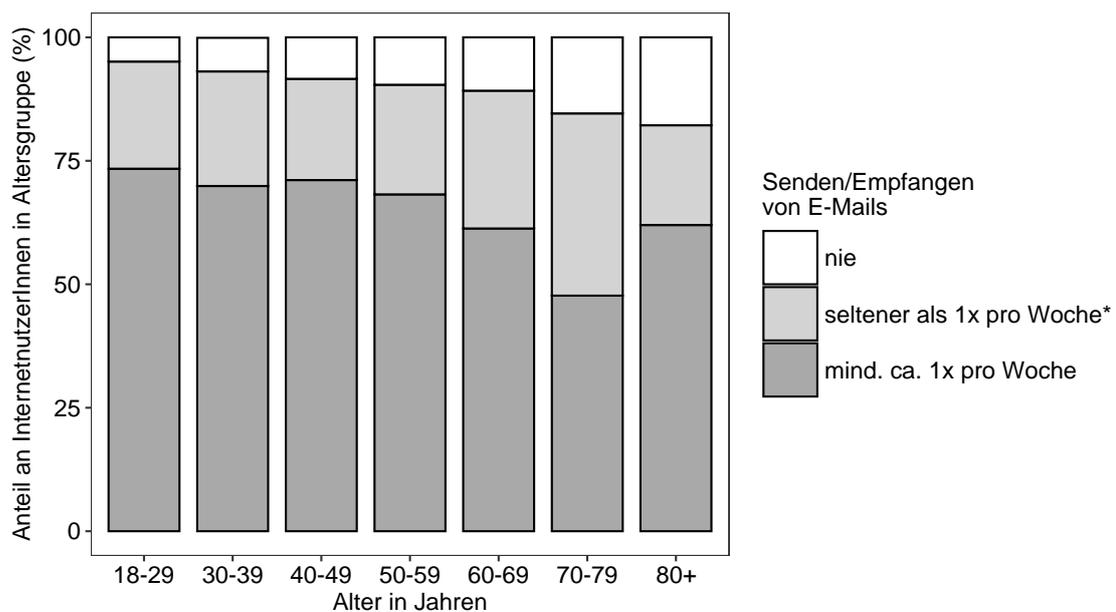
6.2 Nutzung von E-Mail-Diensten

Soll die Stichprobe für die Befragung über eine Liste von E-Mail-Adressen rekrutiert bzw. kontaktiert werden, so ist es erforderlich, dass alle zur Zielpopulation gehörenden Personen neben dem allgemeinen Zugang zum Internet über eine E-Mail-Adresse verfügen und mit dieser Form

¹² Der Wortlaut der Fragestellung weicht in den Jahren 2012 und 2014 gerüfungfügig ab und lautet „Nutzen Sie privat das Internet?“.

der Kommunikation vertraut sind und diese auch nutzen. Während das bloße Vorhandensein einer E-Mail-Adresse häufig bereits durch den verwendeten Internetanbieter sichergestellt ist, kann eine aktive und regelmäßige Nutzung von E-Mail-Diensten in der gesamten Bevölkerung nicht vorausgesetzt werden.

Das Statistische Bundesamt gibt den Anteil der Personen, die zwar das Internet genutzt, jedoch innerhalb der letzten drei Monate keine E-Mails versendet oder empfangen haben, für die gesamte Bevölkerung mit 10 % an. Unterschiede existieren je nach Bildungsstand. Während der Anteil unter Personen mit hoher Bildung lediglich 3 % beträgt, schätzt das Statistische Bundesamt den Anteil unter gering gebildeten Personen auf 15 % (Destatis 2021b). Die durch die VuMA Touchpoints geschätzte Häufigkeit der Nutzung von E-Mail-Diensten unter Internetnutzerinnen und Internetnutzern ist in Abbildung 5 dargestellt. Betrachtet man zunächst die Kategorie „nie“, so ist auch hier deutlich erkennbar, dass es nicht nur in allen Altersgruppen einen Anteil an Internetnutzerinnen und Internetnutzern gibt, die über keine E-Mail-Adresse erreicht werden können, sondern dass dieser Anteil mit zunehmendem Alter steigt. Es sollte daher beachtet werden, dass ein Rückgriff auf E-Mail-Listen die Zahl erreichbarer Personen selbst innerhalb der Gruppe der Internetnutzerinnen und Internetnutzer noch weiter einschränkt und dass diese Ausfälle vermutlich systematisch nach Alter und Bildungsstand der Befragten erfolgen.



*Die Kategorie wurde durch Zusammenfassen der Antworten "mehrmals pro Monat", "ca. 1x pro Monat" und "seltener" gebildet

Abbildung 5: Häufigkeit der Nutzung von E-Mail-Diensten innerhalb der Gruppe von Internetnutzerinnen und Internetnutzern, Datenbasis: VuMA Touchpoints 2021 (VuMa 2021)

Neben der allgemeinen Nutzung eines E-Mail-Postfaches muss gewährleistet sein, dass die Einladung bzw. Aufforderung zur Teilnahme an der Befragung die angeschriebenen Personen zeitnah erreicht. Nutzen Befragte ihr E-Mail-Konto nur in unregelmäßigen Abständen, so kann eine rechtzeitige Teilnahme ggf. gefährdet sein. Selbst in den unteren Altersgruppen verwenden jedoch nur weniger als 75 % der Internetnutzerinnen und Internetnutzer mindestens 1 Mal pro

Woche E-Mail-Dienste, auch hier sinkt die Aktivität mit steigendem Alter. Wird bereits nach einer Woche ein Erinnerungsschreiben versendet oder die Datenerhebung gar nach einem kurzen Zeitraum bereits gestoppt, so muss damit gerechnet werden, dass ein Teil der Befragten die Aufforderung zur Teilnahme noch gar nicht gelesen hat.

6.3 Nutzung von Smartphones

Durch die steigende Popularität von Onlinebefragungen bei gleichzeitiger Verbreitung von mobilen Endgeräten rücken Befragungen, die für die Beantwortung auf Smartphones optimiert sind, immer weiter in den Fokus. Eine Teilnahme an einer solchen Befragung ist nur dann möglich, wenn die Befragten Zugang zu einem mobilen Endgerät haben und in der Lage sind, dieses zum Zweck der Durchführung einer Befragung zu bedienen.

Da zumindest bei ALLBUS und VuMA Touchpoints explizit darauf hingewiesen wird, dass die Nutzung des Internets die Verwendung aller internetfähigen Geräte einschließt, ist davon auszugehen, dass der Anteil an Nutzerinnen und Nutzern von Smartphones nicht den Anteil an Internetnutzerinnen und Internetnutzern insgesamt übersteigt. In den VuMA Touchpoints wird die Nutzung von Smartphones über eine Mehrfachauswahl genutzter internetfähiger Endgeräte erhoben, welche jedoch nur bei positiver Beantwortung der Frage zur Internetnutzung erreicht wird.¹³ Da es theoretisch möglich ist, dass einzelne Befragte trotz der Nutzung eines Smartphones durch die Filterführung gar nicht erst zur Frage gelangen, stellen die dargestellten Werte möglicherweise konservative Schätzungen dar.¹⁴ Der Anteil an Smartphone Nutzerinnen und -nutzern an der Gesamtbevölkerung ist in Abbildung 6 dargestellt; vor allem in den älteren Altersgruppen zeigt sich, dass die Nutzung eines Smartphones auch unter Internetnutzerinnen und Internetnutzern keinesfalls selbstverständlich ist. Der auf Basis der VuMA Touchpoints geschätzte Anteil von Smartphone Nutzerinnen und -nutzern in der Altersgruppe der 70- bis 80-Jährigen beträgt lediglich etwa 30 %, unter den Über-80-Jährigen sogar nur etwa 10 %. Die Zahlen des Statistischen Bundesamtes für das Jahr 2020 liegen etwas höher: für die Altersgruppe der Über-64-Jährigen belaufen sich die Schätzungen für eine Smartphone Nutzung hier auf 74 % der Internetnutzerinnen und Internetnutzer, dies entspräche 48 % der Gesamtbevölkerung in dieser Altersgruppe.¹⁵ Die Daten der VuMA Touchpoints zeigen jedoch deutlich, dass hier hinsichtlich der Smartphone Nutzung eine sehr heterogene Gruppe zusammengefasst wird.

¹³ Die entsprechende Frage lautet: „Hier sind einige Geräte aufgeführt. Sagen Sie mir bitte, ob Sie mit diesen in den letzten 12 Monaten das Internet genutzt haben, egal ob zu Hause, am Arbeitsplatz, in der Schule, Universität, an einem anderen Ort oder unterwegs (...Smartphone oder Handy, egal, ob per Browser oder App)“.

¹⁴ Es könnte jedoch auch argumentiert werden, dass Personen, die bei der Frage nach der Nutzung des Internets ihr eventuell vorhandenes Smartphone nicht einbeziehen, vermutlich nicht zum Personenkreis gehören, welcher bereit und in der Lage ist, eine Onlinebefragung mit dem Smartphone durchzuführen.

¹⁵ Der Anteil an der Gesamtbevölkerung wurde aus den Daten der Publikation der IKT selbst berechnet, daher können Rundungsfehler auftreten.

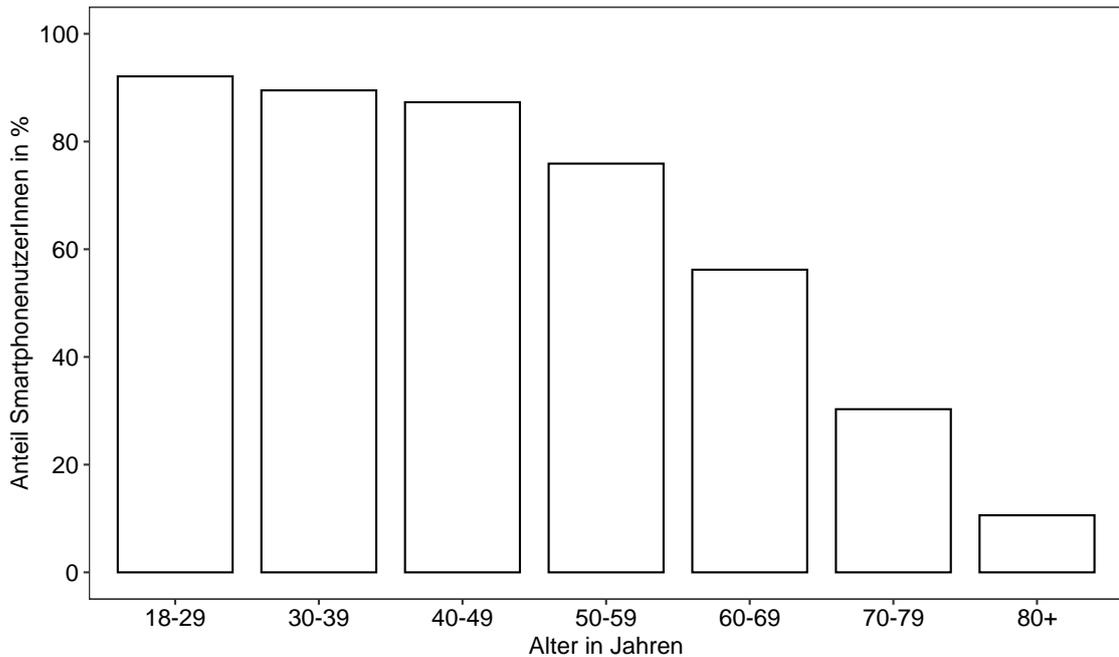


Abbildung 6: Smartphonennutzung nach Alter in Deutschland, Quelle: VuMA Touchpoints 2021 (VuMa 2021)

6.4 Untersuchung des Zusammenhangs zwischen soziodemografischen Merkmalen und der Internetnutzung

Da Personen höheren Alters nicht mit einer selbstverständlichen Nutzung des Internets aufgewachsen sind und beim Erwerb digitaler Kompetenzen häufig auf Selbstaneignung angewiesen sind, können gerade innerhalb der älteren Bevölkerung erhebliche Unterschiede bei der Nutzung des Internets nach soziodemografischen Charakteristika wie Bildung, Geschlecht oder gesundheitlicher Situation erwartet werden. Zur Untersuchung solcher Differenzen innerhalb der älteren Bevölkerung eignen sich insbesondere die Daten der SHARE-Studie, in welcher pro Erhebungswelle mehr als 4.400 Personen in dieser Altersgruppe befragt wurden. Die Auswahl der Befragten erfolgt auf Basis einer Einwohnermeldeamtsstichprobe und enthält neben soziodemografischen Merkmalen auch Informationen zum Gesundheitszustand der Befragten. Trotz der Einschränkung, dass die aktuellsten Daten für 2015 vorliegen, erscheinen die Daten bei SHARE aufgrund ihrer hohen Qualität, des hohen Stichprobenumfangs und der Verfügbarkeit von Gesundheitsmerkmalen geeignet, um mögliche systematische Unterschiede in der Internetnutzung innerhalb der Altersgruppen zu untersuchen. Beachtet werden muss hierbei, dass bei SHARE eine Internetnutzung innerhalb der letzten sieben Tage abgefragt wird, weshalb entsprechende Schätzungen ggf. etwas geringer ausfallen als bei der Erhebung der grundsätzlichen Nutzung des Internets.

Wie Abbildung 7 entnommen werden kann, weisen die Daten der SHARE-Studie auf einen deutlichen Interaktionseffekt zwischen Alter und Bildung hinsichtlich der Nutzung des Internets hin.¹⁶ Ein starker Effekt der Bildung zeigt sich bereits bei den 50- bis-60-Jährigen, die Kluft vergrößert sich mit zunehmendem Alter jedoch erheblich. Zwar weist ALLBUS einen in der betrachteten Altersgruppe deutlich geringeren Stichprobenumfang auf, die gleiche Tendenz ist jedoch auch in den ALLBUS-Daten erkennbar.¹⁷ Die Tatsache, dass der Effekt der Bildung nicht nur in den höchsten Altersgruppen vorhanden ist, legt nahe, dass bei Onlinebefragungen zumindest in den nächsten Jahren mit einer systematischen Unterrepräsentation von Personen mit niedriger Bildung gerechnet werden muss. Derzeit scheint das Internet im sehr hohen Alter nahezu ausschließlich von Personen mit mittlerer und hoher Bildung genutzt zu werden.

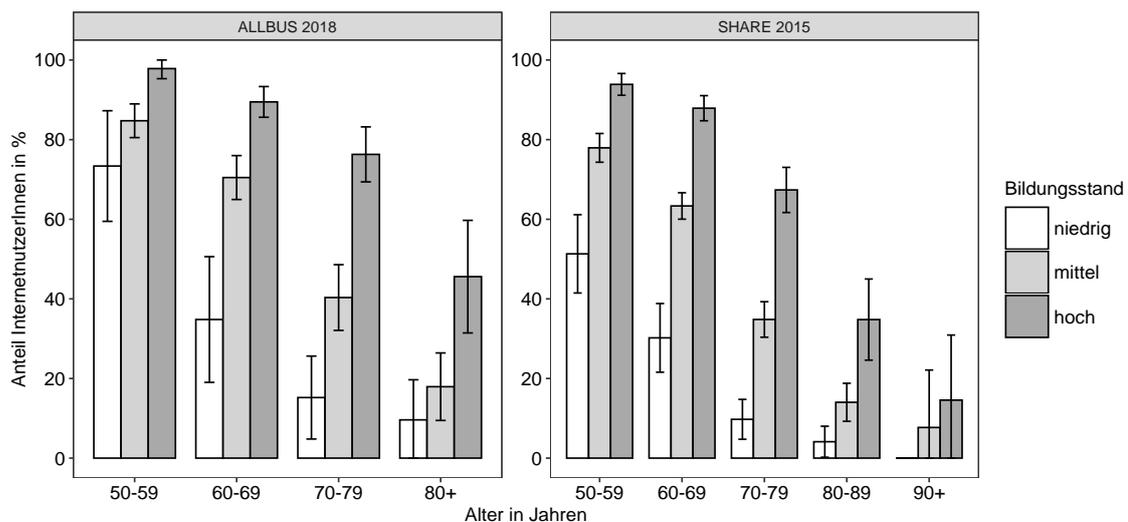


Abbildung 7: Internetnutzung nach Bildungsstand und Alter, Datenbasis: ALLBUS 2018 (GESIS 2019) und SHARE Welle 6 (2015) (Börsch-Supan 2019)

Ebenfalls zeigt sich in den Datensätzen der SHARE-Studie ein im Alter zunehmender Effekt des Geschlechts auf die Internetnutzung. Wie Abbildung 8 entnommen werden kann, nutzen Frauen ab 60 Jahren deutlich seltener das Internet als Männer in der gleichen Altersklasse. Zwar muss erneut einschränkend angemerkt werden, dass die Daten bereits aus dem Jahr 2015 stammen. Die erheblichen Differenzen in der Internetnutzung bereits im Alter ab 65 Jahren lassen jedoch vermuten, dass auch in den kommenden Jahren zumindest in den höheren Altersgruppen noch mit einer deutlich geringeren Internetnutzung unter Frauen gerechnet werden muss.

¹⁶ Die Einteilung in drei Bildungsklassen erfolgte hier anhand der ISCED 97-Klassifikation (BMBF 2013), wobei die Stufen 0 bis 2 zu niedriger Bildung, 3 und 4 zu mittlerer Bildung und 5 und 6 zu hoher Bildung zusammengefasst wurden.

¹⁷ Der höhere Anteil an Internetnutzerinnen und Internetnutzern bei ALLBUS vor allem unter den Personen mit geringer Bildung in der Altersgruppe der 50- bis 59-Jährigen könnte einerseits auf die unterschiedliche Fragestellung, andererseits auf den späteren Erhebungszeitraum zurückzuführen sein.

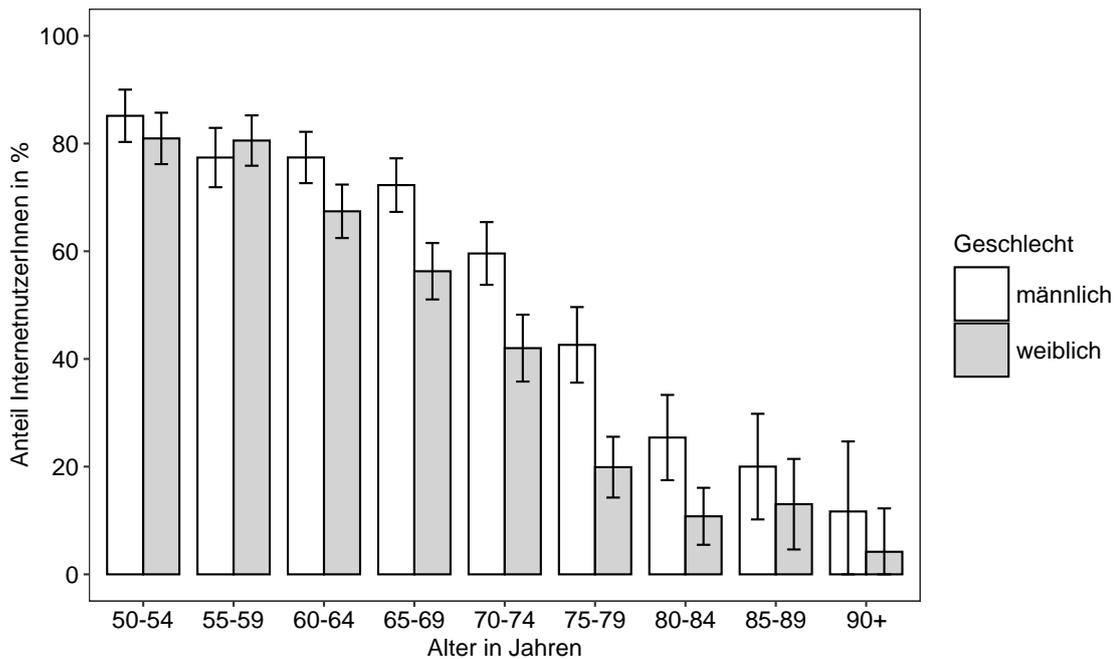


Abbildung 8: Internetnutzung nach Geschlecht und Alter, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)

Bei Befragungen von Patientinnen und Patienten ist zu erwarten, dass je nach Krankheitsbild der Gesundheitszustand bzw. der funktionelle Status der Befragten von der Allgemeinbevölkerung abweicht. Die Frage, ob sich der Gesundheitszustand der Befragten auf die Teilnahmewahrscheinlichkeit bei Onlinebefragungen auswirken kann, ist in diesem Kontext daher von besonderem Interesse. In der SHARE-Studie wird diesbezüglich erhoben, ob die Befragten innerhalb der letzten 6 Monate aus gesundheitlichen Gründen in ihren alltäglichen Verrichtungen eingeschränkt waren.¹⁸ Dabei zeigt sich, dass funktionelle Einschränkungen im Alltag negativ mit der selbstberichteten Internetnutzung assoziiert sind (siehe Abbildung 9). Bei den Über-70-Jährigen nutzten nur etwa halb so viele Personen mit starken Einschränkungen das Internet als dies bei Personen ohne Einschränkungen im Alltag der Fall war. Die wenig eindeutigen Befunde in der Gruppe der Über-90-Jährigen sind mit hoher Wahrscheinlichkeit auf geringe Fallzahlen zurückzuführen.

¹⁸ Die Frage lautet: „In welchem Maße waren Sie mindestens während der letzten sechs Monate wegen eines gesundheitlichen Problems bei alltäglichen Verrichtungen eingeschränkt? (Stark eingeschränkt / Eingeschränkt, aber nicht stark / Nicht eingeschränkt)“

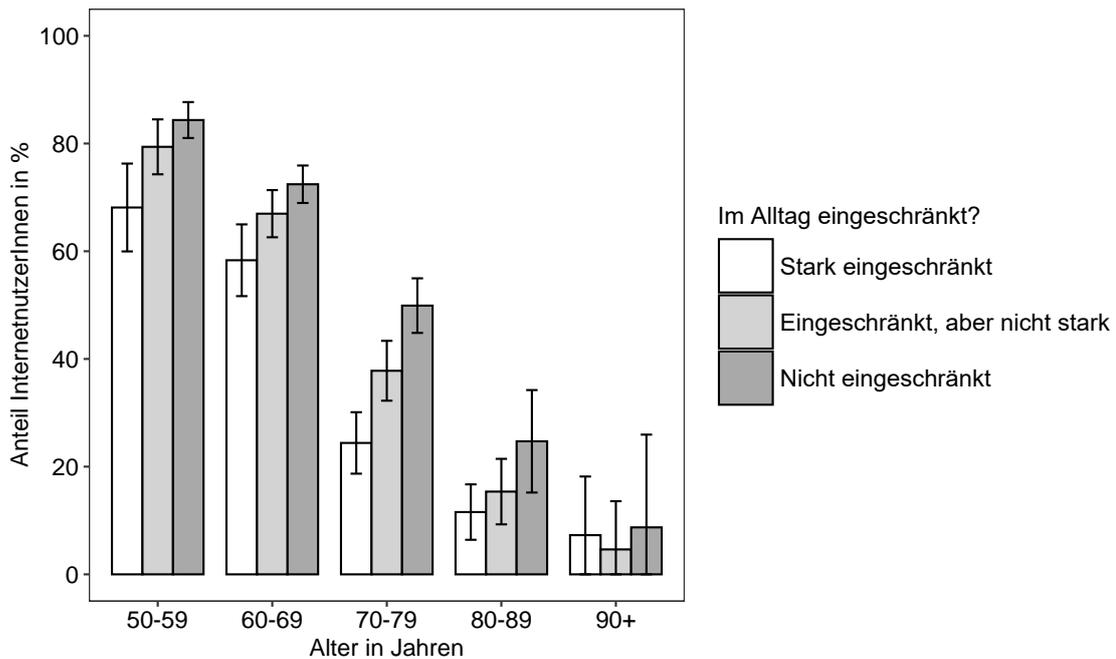


Abbildung 9: Internetnutzung in Abhängigkeit von gesundheitlichen Einschränkungen im Alltag, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)

Ein ähnliches Bild ergibt sich, wenn die Summe der ärztlich diagnostizierten Krankheiten innerhalb der jeweiligen Altersgruppen als Indikator für den Gesundheitszustand betrachtet wird.¹⁹ Dargestellt ist dies in Abbildung 10, kategorisiert nach der Anzahl an selbstberichteten diagnostizierten Erkrankungen (Kategorie 1: Personen ohne diagnostizierte Krankheiten; Kategorie 2: eine oder zwei diagnostizierte Krankheiten; Kategorie 3: drei oder mehr diagnostizierte Krankheiten). Der Zusammenhang zwischen dem Gesundheitszustand und der Nutzung des Internets ist vor allem bei den Personen über 70 Jahren deutlich zu erkennen, besteht jedoch auch in den jüngeren Altersgruppen.

¹⁹ Erfasst werden bei SHARE folgende Krankheiten: Herzinfarkt einschließlich Myokardinfarkt, Koronarthrombose oder andere Herzkrankheiten einschließlich Herzinsuffizienz; Bluthochdruck; hohe Cholesterinwerte; Schlaganfall einschließlich Durchblutungsstörungen im Gehirn; Diabetes oder hohe Blutzuckerwerte; Chronische Erkrankungen der Lunge wie chronische Bronchitis oder Lungenemphysem; Krebs oder bösartige Tumore, einschließlich Leukämie und Lymphdrüsenkrebs ausschließlich kleinerer Hautkrebsarten; Magengeschwür, Zwölffingerdarmgeschwür; Parkinson-Krankheit; Grauer Star; Oberschenkelhalsbruch oder Hüftfraktur; andere Brüche oder Frakturen; Alzheimer, Demenz, Senilität oder andere schwere Gedächtnisstörung; andere seelische oder emotionale Störungen wie Angststörungen, Nervenleiden oder psychische Probleme; rheumatische Gelenkentzündung, Arthritis; Arthrose oder andere Formen von Rheuma; chronische Nierenkrankheit; andere Krankheiten

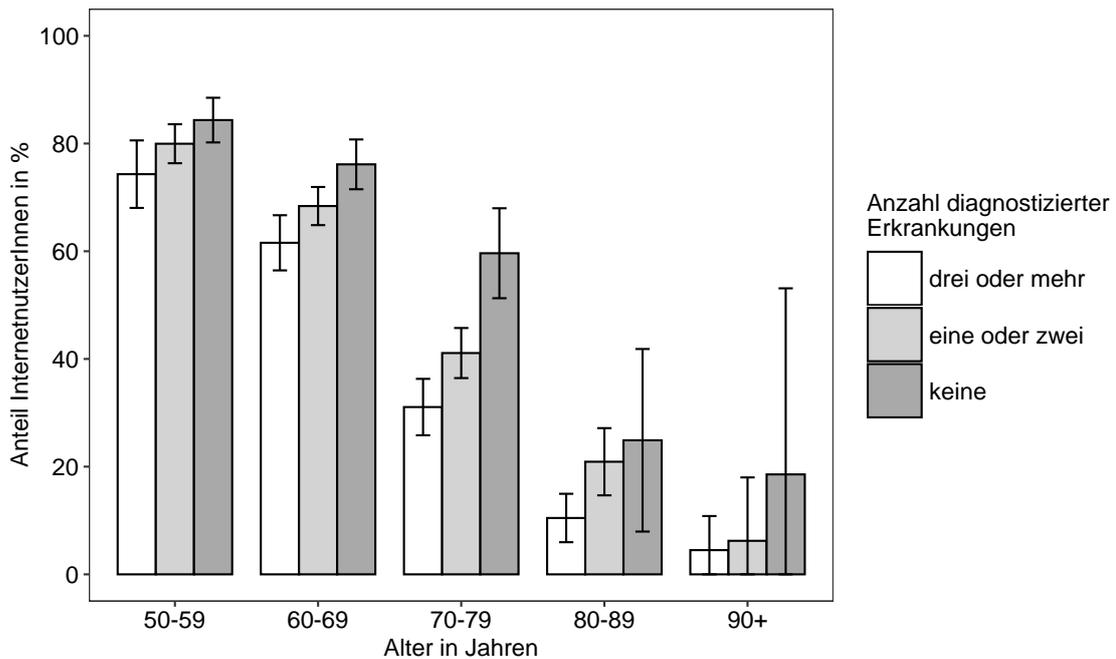


Abbildung 10: Internetnutzung in Abhängigkeit von der Summe der diagnostizierten Erkrankungen, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)

6.5 Multivariable Betrachtung potenzieller Prädiktoren für die Nutzung des Internets

Um zu untersuchen, ob die gezeigten Effekte von Alter, Geschlecht, Bildung und Gesundheitsstatus bestehen bleiben, wenn für die jeweils anderen Merkmaleffekte adjustiert wird, wurde die Wahrscheinlichkeit zur Nutzung des Internets in einem binomial-logistischen Regressionsmodell auf Basis der SHARE-Daten geschätzt. Da unterschiedliche Variablen als Indikatoren für den Gesundheitsstatus herangezogen werden können (funktionelle Einschränkungen im Alltag, Anzahl von Erkrankungen bzw. Multimorbidität, Vorhandensein bestimmter diagnostizierter Erkrankungen) und hohe Korrelationen zwischen diesen Variablen zu erwarten sind, wurden drei separate Regressionsmodelle berechnet.²⁰ Dabei wurden die Variablen Alter, Geschlecht und Bildung in jedes Modell aufgenommen, während die jeweiligen Indikatorvariablen für den Gesundheitsstatus variierten (vgl. Abbildung 11, Abbildung 12 und Abbildung 13). Die Variable „Alter“ wurde dabei in 10-Jahres-Schritten kategorial in die Modelle aufgenommen, um mit den vorherigen Darstellungen in diesem Kapitel konsistent zu sein.

Im ersten Regressionsmodell wurde das selbstberichtete Ausmaß an funktionellen Einschränkungen im Alltag (Stark eingeschränkt / Eingeschränkt / Nicht eingeschränkt) als Indikator für den Gesundheitsstatus miteinbezogen. In Abbildung 11 sind die geschätzten Koeffizienten des Modells mit den eingeschlossenen Prädiktoren abgebildet. Dargestellt sind die Odds-Ratio sowie die zugehörigen 95 %-Konfidenzintervalle (KI).

²⁰ Wenn zwei oder mehrere Prädiktoren sehr stark miteinander assoziiert sind, kann die Interpretation der Ergebnisse des Regressionsmodells erschwert werden.

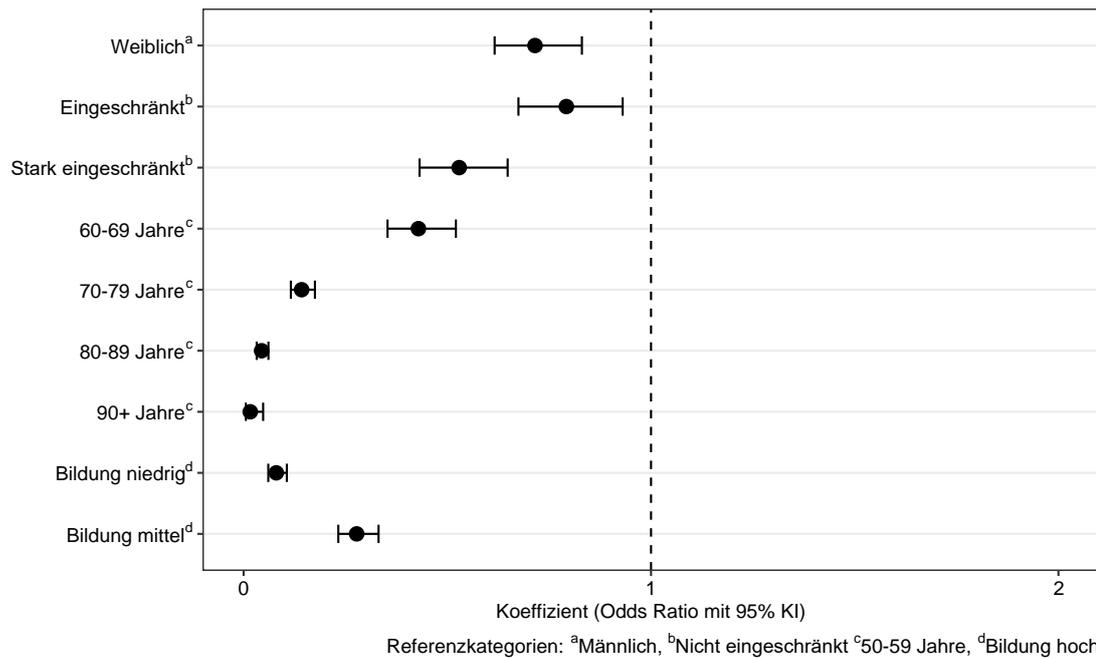


Abbildung 11: Binomial-logistische Regression zu Internetnutzung in den letzten sieben Tagen mit Einschränkungen im Alltag als Indikator für den Gesundheitsstatus, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)

Es zeigt sich, dass auch bei multivariabler Betrachtung der potenziellen Prädiktoren für die Nutzung des Internets alle patientenseitigen Merkmale einen statistisch signifikanten Effekt in die erwartete Richtung aufweisen.

Im zweiten Regressionsmodell wurde die Summe der diagnostizierten Erkrankungen (keine Erkrankung / eine oder zwei Erkrankungen / drei oder mehr Erkrankungen) als Indikator für den Gesundheitsstatus miteinbezogen. In Abbildung 12 zeigt sich, dass auch in diesem Modell alle untersuchten Merkmale einen statistisch signifikanten Effekt in die erwartete Richtung aufweisen. Die Effekte der Kategorien „eine oder zwei Erkrankungen“ bzw. „drei oder mehr Erkrankungen“ sind dabei vergleichbar mit den Effekten, die im ersten Regressionsmodell für die Kategorien „Eingeschränkt“ und „Stark eingeschränkt“ berechnet wurden.

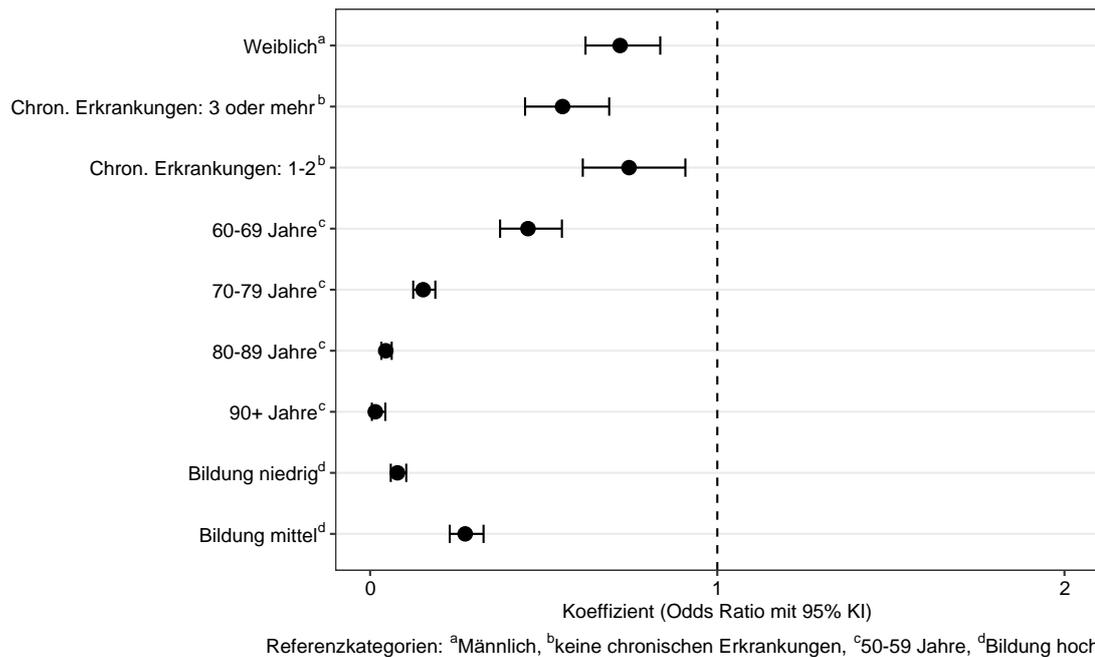


Abbildung 12: Binomial-logistische Regression zu Internetnutzung in den letzten sieben Tagen mit Summe der diagnostizierten Erkrankungen als Indikator für den Gesundheitsstatus, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)

In einem dritten Modell wurde die Internetnutzung in Abhängigkeit einzelner Erkrankungen bzw. Krankheitsgruppen²¹ betrachtet, bei gleichzeitigem Einschluss der patientenseitigen Variablen Alter, Geschlecht und Bildung. Die Ergebnisse sind grafisch in Abbildung 13 dargestellt. Es zeigt sich, dass – im Gegensatz zur Summe unterschiedlicher Erkrankungen – die meisten Krankheiten, wenn deren bloßes Vorhandensein (Ja / Nein) in das Modell einzeln aufgenommen wird, für sich allein keinen bedeutenden Einfluss auf die Internetnutzung haben. Lediglich für neurologische Erkrankungen sowie für Krebserkrankungen konnte ein statistisch signifikantes Ergebnis gezeigt werden.

²¹ Herzinfarkt, Bluthochdruck und Schlaganfall wurden in der Kategorie „Kardiovaskuläre Erkrankungen“ zusammengefasst; die Parkinson-Krankheit wurde gemeinsam mit Alzheimer, Demenz, Senilität oder andere schwere Gedächtnisstörung in der Kategorie „Neurologische Erkrankung“ zusammengefasst; rheumatische Gelenkentzündung, Arthritis wurde mit Arthrose oder andere Formen von Rheuma unter „Rheumatische Erkrankung“ zusammengefasst; alle weiteren bei SHARE erfassten Erkrankungen, die in der Grafik nicht im Einzelnen dargestellt sind, wurden in der Kategorie „Andere Krankheit“ zusammengefasst.

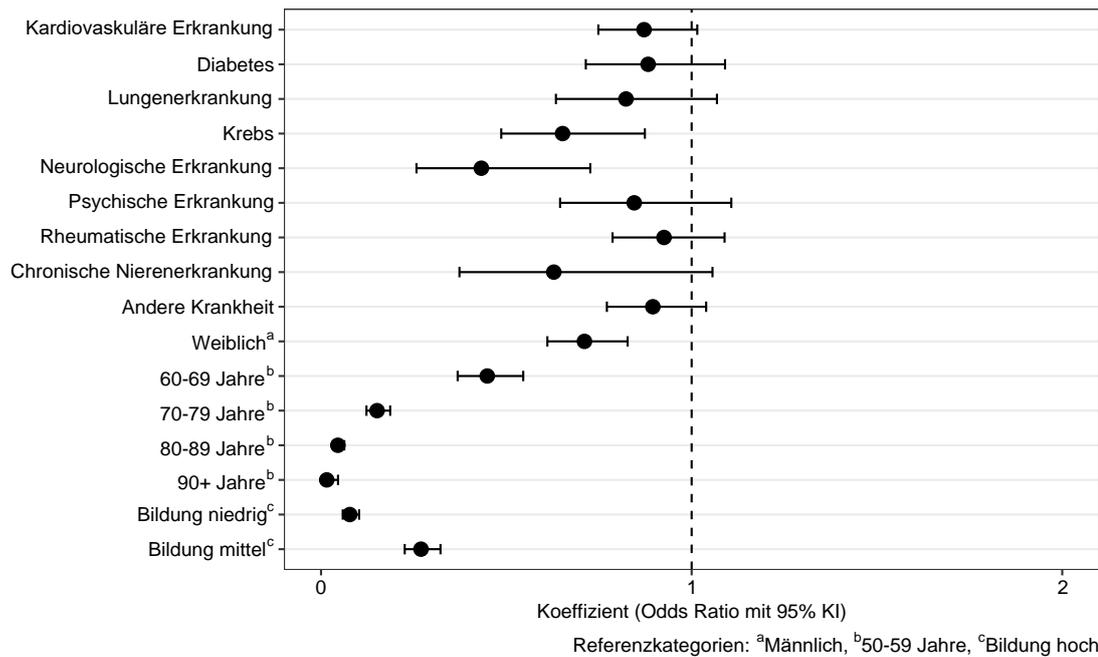


Abbildung 13: Binomial-logistische Regression zu Internetnutzung in den letzten sieben Tagen mit einzelnen Erkrankungen bzw. Erkrankungsgruppen als Indikatoren für den Gesundheitsstatus, Datenbasis: SHARE Welle 6 (2015) (Börsch-Supan 2019)

Zusammengefasst wird aufgrund der in diesem Abschnitt durchgeführten Analysen deutlich, dass die zuvor detektierten Effekte von Alter, Geschlecht, Bildung und Gesundheitsstatus auf die Internetnutzung bestehen bleiben, wenn für die jeweils anderen Merkmalseffekte adjustiert wird. Dies lässt beispielsweise den Schluss zu, dass ein hohes Maß an diagnostizierten Erkrankungen bzw. Einschränkungen im Alltag nicht nur deshalb negativ mit der Nutzung des Internets assoziiert ist, weil sich unter diesen Patientinnen und Patienten sehr viele ältere Personen befinden, sondern dass der Gesundheitszustand per se signifikant mit der Internetnutzung assoziiert zu sein scheint. Eine Ausnahme stellen diesbezüglich das Vorhandensein bzw. Nicht-Vorhandensein der meisten Erkrankungen dar, wenn diese als separate Prädiktoren einzeln in das Modell aufgenommen wurden.

7 Möglichkeiten der Umsetzung einer Onlinebefragung mit Patientinnen und Patienten

Onlinebefragungen werden in den letzten Jahren immer häufiger eingesetzt. So wird mittlerweile ein großer Teil der standardisierten Befragungen der Institute des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e. V. (ADM) als Onlinebefragung durchgeführt (ADM 2020). Die vorrangigen Vorteile von Onlinebefragungen liegen in ihrer vergleichsweise kostengünstigen Umsetzung und ihrer Fehlerminimierung bei der Beantwortung des Fragebogens sowie den flexiblen Darstellungsmöglichkeiten (z. B. Einbindung von Multimediaelementen im Fragebogen) und der schnellen und effizienten Überführung der Antworten in einen Datensatz (Callegaro et al. 2015, Bandilla 2015). Dennoch wurden andere selbstadministrierte Erhebungsmodi, insbesondere postalische Paper-Pencil-Befragungen, bisher nicht verdrängt. Dies hat vielseitige Gründe und liegt u. a. daran, dass oftmals vollständige Listen mit Kontaktinformationen (z. B. E-Mail-Listen) der Zielpopulationen nicht immer verfügbar sind (Sampling-Probleme) oder dass nicht alle Mitglieder einer Zielpopulation über einen Internetzugang verfügen (Coverage-Probleme, siehe auch Kapitel 6) (Schonlau und Couper 2017).

Bei der Entscheidung, welcher Erhebungsmodus am besten mit den Zielen und Rahmenbedingungen einer bestimmten Befragung vereinbar ist, gibt es eine Vielzahl unterschiedlicher Faktoren zu beachten. Das folgende Kapitel gibt einen Überblick über die Vor- und Nachteile von onlinebasierten Befragungen. Dabei sollen unterschiedliche Formen von Onlinebefragungen dargestellt und herkömmlichen papierbasierten Befragungen gegenübergestellt werden. Zudem soll auf Besonderheiten hinsichtlich Design und Fragebogenkonstruktion und der Kontaktierung von Befragten sowie auf die spezifischen IT-technischen Voraussetzungen und rechtlichen Rahmenbedingungen bei der Umsetzung von Onlinebefragungen, insbesondere im Kontext der gesetzlichen Qualitätssicherung, eingegangen werden.

In diesem Kapitel wird ausschließlich auf selbstadministrierte Befragungsformen eingegangen, da andere Formen der Befragung, wie etwa Face-to-Face-Interviews oder assistierte Computerumfragen für Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung zum aktuellen Zeitpunkt nicht vorgesehen sind und es sich dementsprechend auch bei den bisher entwickelten Patientenbefragungen ausschließlich um selbstadministrierte Befragungen handelt (IQTIG 2018b, IQTIG 2018a, IQTIG 2020).

7.1 Formen von Onlinebefragungen

Onlinebefragungen erlauben ein flexibles Vorgehen hinsichtlich unterschiedlicher Aspekte der Umsetzung (z. B. Weg der Kontaktierung, notwendige Software und Hardware zur Beantwortung, Design und Darstellung der Items im Fragebogen) und zeichnen sich daher durch ein hohes Maß an Diversität aus. Grundsätzlich kann zwischen onlinebasierten Single-Mode- und Mixed-Mode-Befragungen unterschieden werden (Schnell 2019). Eine Single-Mode-Onlinebefragung ist dadurch gekennzeichnet, dass die Befragung in nur einem Modus durchgeführt werden kann,

d. h. nur online. Bei Mixed-Mode-Befragungen kann der Fragebogen hingegen in unterschiedlichen Modi beantwortet werden, z. B. Online- und Papierfragebögen, in der Regel je nach individueller Präferenz der Befragten. An dieser Stelle ist festzuhalten, dass sich die Begrifflichkeiten „Single-Mode“ und „Mixed-Mode“ in diesem Bericht ausschließlich auf die Form der Datenerhebung bezieht und nicht auf die Art der Kontaktierung²² (Callegaro et al. 2015).

In den folgenden Abschnitten werden unterschiedliche Formen von Single-Mode- und Mixed-Mode-Befragungen vorgestellt und ihre Vor- und Nachteile postalischen Befragungen gegenübergestellt.

7.1.1 Single-Mode

Bei Single-Mode-Onlinebefragungen (auch „Web Survey“ genannt) erhalten die Befragten in der Regel einen Link und / oder QR-Code, mit dem sie zum Fragebogen weitergeleitet werden. Bei postalischen Paper-Pencil-Befragungen im Single-Mode, wie dies für die bisher entwickelten Patientenbefragungen im Rahmen der externen Qualitätssicherung zum aktuellen Stand vorgesehen ist, werden die vorliegenden Wohnadressen verwendet, um den Testpersonen einen Fragebogen per Post zuzuschicken. Bei Onlinebefragungen gibt es dagegen auch verschiedene andere Möglichkeiten der Rekrutierung bzw. Kontaktierung der Testpersonen. So können Einladung und Link zur Teilnahme an der Onlinebefragung per E-Mail, SMS o. ä. verschickt werden oder aber als postalisches Anschreiben, das den Link (bzw. QR-Code) und die notwendigen Zugangsdaten enthält.

Um sicherzustellen, dass der Fragebogen pro Person nur einmal ausgefüllt wird, bekommen die Befragten üblicherweise eine individuelle Kennung, z. B. anhand eines individuellen Passworts. Dadurch kann den Teilnehmerinnen und Teilnehmern auch die Möglichkeit gegeben werden, die Befragung jederzeit zu unterbrechen und später ohne Verlust der bisher eingegebenen Antworten fortzusetzen (Couper 2008).

Weitere Formen von Single-Mode-Onlinebefragungen

Weitere Formen von Single-Mode-Onlinebefragungen sind E-Mail-Surveys, Intercept-Surveys, Mobile-Web-Surveys sowie Befragungen mit elektronischen Eingabegeräten wie z. B. Tablets, welche in Einrichtungen verteilt werden. Diese Arten von Onlinebefragungen sind bei der Umsetzung mit verschiedenen technischen bzw. methodischen Anforderungen verbunden, die im Kontext der gesetzlichen Qualitätssicherung nur eingeschränkt oder gar nicht realisiert werden können und daher hier nicht im Fokus stehen. Für einen umfassenden Überblick erfolgt dennoch eine kurze Darstellung.

Bei E-Mail-Surveys bekommen die Befragten einen Fragebogen per E-Mail zugeschickt, z. B. als PDF oder als Formular in der E-Mail selbst und schicken diesen per E-Mail auch wieder zurück. Diese Art der Onlinebefragung erfordert standardisierte E-Mail-Programme und Eingabegeräte

²² Der Begriff „Mixed-Mode“ wird in der Literatur unterschiedlich verwendet. Manche Autorinnen und Autoren verwenden den Begriff, wenn die Kontaktierung in einem anderen Modus (z. B. postalisch als Brief) erfolgt als die Art der Datenerhebung (z. B. als Onlinefragebogen), oder wenn die erste Erhebungswelle einer Umfrage in einem anderen Modus erfolgt als die zweite (Medway und Fulton 2012).

(Schnell et al. 2018) und ist für Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung nicht geeignet, da die technischen Rahmenbedingungen nicht gegeben sind. Bei Intercept-Surveys werden per Zufallsauswahl beliebige Besucherinnen bzw. Besucher einer Webseite dazu eingeladen, an einer Befragung teilzunehmen. Hierbei handelt es sich nicht um eine echte Zufallsstichprobe, da es sich um eine selbstrekrutierte Stichprobe handelt und es somit nicht möglich ist, Aussagen über eine Grundgesamtheit zu tätigen und Verallgemeinerungen zu machen (Schnell et al. 2018). Mobile-Web-Surveys sind Onlinebefragungen, bei denen der Fragebogen über eine App auf einem Smartphone oder Tablet ausgefüllt wird (Couper et al. 2017). Die Umsetzung von Mobile-Web-Surveys ist technisch aufwendig und eignet sich nicht für alle Zielgruppen. So müssen die Befragten z. B. eine App installieren, die erst programmiert werden müsste, um an der Befragung teilzunehmen. Zudem ist nicht davon auszugehen, dass alle Personen der Zielgruppe von Patientenbefragungen über ein Smartphone und / oder ein Tablet verfügen (siehe Kapitel 6.3).

Bei Befragungen über Tablets oder PCs, die den Einrichtungen gehören, müssen von Mitarbeiterinnen und Mitarbeitern der Einrichtungen diese übergeben und nach der Beantwortung der Fragebögen wieder eingesammelt werden. Dadurch kann nicht gewährleistet werden, dass die Beantwortung des Fragebogens durch das medizinische Personal unbeeinflusst bleibt. Zudem lassen sich solche Befragungen nur mit Patientinnen und Patienten durchführen, die sich zum Zeitpunkt der Befragung in der Einrichtung befinden, weswegen diese Methode bei Qualitätssicherungsverfahren im ambulanten Bereich oder bei Verfahren, im Rahmen derer Patientinnen und Patienten erst nach Abschluss ihrer Behandlung befragt werden, ungeeignet ist. Ferner müsste die technische Infrastruktur für diese Art der Befragung bei allen Leistungserbringern vorhanden sein, was mit hohen zusätzlichen Kosten verbunden wäre.

Vorteile von Single-Mode-Onlinebefragungen gegenüber postalischen Befragungen

Ein Vorteil von Single-Mode-Onlinebefragungen gegenüber postalischen Befragungen liegt in den vergleichsweise geringen Kosten bei der Durchführung. Eine Single-Mode-Onlinebefragung ist im Vergleich zu postalischen Befragungen kostengünstiger, da keine bzw. je nach Art der Kontaktierung geringere Druck- und Portokosten und keine Kosten für die Dateneingabe anfallen (Callegaro et al. 2015). Onlinebefragungen benötigen im Rahmen der Datenerhebung weniger zeitliche Ressourcen, was vor allem dann der Fall ist, wenn die Kontaktierung der Befragten per E-Mail erfolgt (Callegaro et al. 2015). Darüber hinaus haben Onlinebefragungen den Vorteil unmittelbarer Datenspeicherung (Bandilla 2015). Daten müssen nicht mehr aus einem Papierfragebogen digitalisiert und korrigiert werden, was einen Medienbruch vermeidet und weniger zeitintensiv und fehleranfällig ist. Zudem stellen Onlinebefragungen eine zeit- und ortsunabhängige Befragungsmöglichkeit dar, da sie nicht nur via Computer, sondern auch auf Smartphones oder Tablets ausgefüllt werden können (Callegaro et al. 2015). Hierbei sind mögliche Datenverzerrungen aufgrund unterschiedlicher Eingabegeräte zu beachten, bspw. indem ein Teil eines Schiebereglers auf einem mobilen Endgerät nicht in der vollen Breite bzw. Länge angezeigt wird. Ein weiterer Vorteil von Onlinebefragungen liegt in der Sicherung der Datenqualität. In einem elektronischen Fragebogen können Wertebereiche für Eingabefelder vorgegeben und Eingaben

einer Plausibilitätsprüfung unterzogen werden. Hiermit wird die Eingabe unzulässiger und unlogischer Werte verhindert. So kann man beispielsweise im Feld zur Frage nach dem Alter der Testpersonen Restriktionen setzen, sodass nur positive numerische Werte eingegeben werden können und es kann ein Minimal- und Maximalwert festgelegt werden. Richtet sich die Befragung bspw. ausschließlich an volljährige Personen, könnte man als Minimalwert 18 und als Maximalwert 120 Jahre festlegen.²³ Damit wird verhindert, dass die Testpersonen unlogische Werte wie z. B. „650“ oder „fünfundsechzig“ eingeben. Bei einer postalischen Befragung müssten solche Werte als fehlend deklariert werden.

Durch die Programmierbarkeit von Onlinefragebögen lassen sich neben der unmittelbaren Datenkontrolle auch komplexe Filterführungen verwirklichen (Bandilla 2015). Bei postalischen Befragungen muss bei Filterfragen einer Sprunganweisung gefolgt werden, was je nach Filterführung kognitiv anspruchsvoll ist, da der vertikale Lesefluss unterbrochen wird und man ggf. eine Seite weiter springen muss. Dadurch kommt es immer wieder vor, dass die Sprunganweisung falsch ausgeführt oder ignoriert wird. Bei Onlinebefragungen ist dies nicht der Fall, da man sofort zu der Frage weitergeleitet wird, die man laut Filter beantworten muss. Zudem ist es bei Onlinebefragungen technisch umsetzbar, die Reihenfolge der Fragen zu randomisieren und dadurch Verzerrungen durch Reihenfolgeeffekte zu vermeiden (Bandilla 2015, Callegaro et al. 2015).

Neben dem Aspekt der Datenqualitätskontrolle bieten Onlinebefragungen die Möglichkeit, auf sogenannte Paradata zurückzugreifen, solange im Voraus das Einverständnis der Befragten eingeholt wurde. Solche Paradata sind z. B. Ausfüllzeit bzw. Verweildauer bei einer Frage oder Informationen darüber, ob und wie oft eine Antwort geändert wurde. Diese Informationen sind beispielsweise im Rahmen der Pretestung von Fragebögen relevant, da potenzielle Probleme bei der Beantwortung aufgedeckt werden können. Mithilfe von Paradata lässt sich ebenfalls sogenanntes „Speeding“ aufdecken. Single-Mode-Onlinebefragungen können Teilnehmerinnen und Teilnehmer verleiten, die Fragen sehr schnell zu lesen und zu beantworten, ohne deren Inhalt adäquat verarbeitet zu haben (Conrad et al. 2017). Mithilfe von Paradata kann überprüft werden, ob Speeding vorliegt und ob zu schnell gegebene Antworten einen Einfluss auf die Datenqualität haben (Greszki et al. 2015).

Weiterhin lassen sich in Onlinebefragungen Multimediainhalte wie Bilder, Audio- oder Videodateien in den Fragebogen einbauen, um die Beantwortung zu erleichtern oder Reaktionen zu den Inhalten dieser Dateien zu erfassen (Bandilla 2015). Dies ist vor allem für die Barrierefreiheit von Befragungen ein wichtiger Vorteil, da mithilfe integrierter Audioaufnahmen etwa Fragen vorgelesen werden können.

Zudem haben Onlinebefragungen gegenüber postalischen Befragungen den Vorteil, dass bestimmte Zielgruppen eine höhere Teilnahmebereitschaft zeigen. So gibt es nicht nur Evidenz, dass junge Bevölkerungsgruppen über Onlinebefragungen besser zu erreichen sind (Toepoel und Lugtig 2014) sondern auch, dass Angehörige sexueller Minderheiten eher digitale gegenüber postalischen Befragungen präferieren (Bates et al. 2019).

²³ Die Wahrscheinlichkeit, dass ein Mensch ein Alter von über 120 Jahren erreicht, ist sehr gering.

Nachteile von Single-Mode-Onlinebefragung gegenüber postalischen Befragungen

Wie bereits beschrieben, können Teilnehmerinnen und Teilnehmer von Onlinebefragungen digital oder postalisch kontaktiert werden. Prinzipiell wird empfohlen, bei Single-Mode-Befragungen für Kontaktierung und Befragung den gleichen Modus zu nutzen (Schnell 2019, Groves et al. 2009). Im Fall von Single-Mode-Onlinebefragungen ist dies eine digitale Kontaktaufnahme mit einem Link zur Onlinebefragung. Die digitale Kontaktaufnahme per E-Mail und / oder SMS hat dabei den Nachteil, dass vollständige Listen mit E-Mail-Adressen oder Mobilnummern der Zielpopulation vorliegen müssen, aus denen die Zufallsstichprobe gezogen werden kann. Für Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung sind diese Informationen jedoch im Gegensatz zu Adressdaten nicht automatisch verfügbar (siehe Abschnitt 7.4). Diese Informationen müssten daher von den Leistungserbringern separat erfragt werden, was zusätzlichen Aufwand bedeuten würde. Außerdem verfügen nicht alle Patientinnen und Patienten über ein Handy oder eine E-Mail-Adresse und die Erfassung von E-Mail-Adressen und Handynummern ist fehleranfällig, was zur Nichtzustellbarkeit von Befragungen führen kann. Ein weiteres mögliches Problem der digitalen Kontaktierung ist, dass die E-Mail mit der Onlinebefragung als Spam gewertet werden könnte, was die Beantwortung der Onlinebefragung unwahrscheinlich macht. Dieser Herausforderung kann begegnet werden, indem der Zugang zu Single-Mode-Onlinebefragungen per Brief versendet wird. Studien zeigen, dass eine postalische Kontaktaufnahme bei Onlinebefragungen die Teilnehmerate gegenüber digitaler Kontaktaufnahme sogar erhöhen kann (Dykema et al. 2012, Sakshaug et al. 2018). Ein Grund hierfür könnte sein, dass per Post verschickte Einladungen die Legitimität und Bedeutung der Befragung unterstreichen (Dykema et al. 2012, Dillman et al. 2014a).

Eine weitere Herausforderung von Onlinebefragungen ist es, sie so zu programmieren, dass die Darstellung unabhängig vom Eingabegerät oder der Software (z. B. Browser) möglichst identisch ist, da Onlinebefragungen üblicherweise nicht ausschließlich am PC ausgefüllt werden, sondern auch am Smartphone oder Tablet. Diese Möglichkeit muss entsprechend bei der Gestaltung bzw. Programmierung des Fragebogens beachtet werden. In dieser Hinsicht ist auch sicherzustellen, dass die Onlinebefragung möglichst unabhängig von der Internetverbindung schnell geladen werden kann, da es sonst wahrscheinlich ist, dass die Teilnahmemotivation der Befragten sinkt (Schnell 2019). Bei Onlinebefragungen ist zudem zu beachten, den Befragungsaufwand zu minimieren, indem bspw. vermieden wird, dass die Befragten eine besondere App installieren müssen, da dies die Teilnahmebereitschaft ebenfalls senken kann (Wenz et al. 2019).

Ein weiterer, für Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung relevanter Nachteil von Single-Mode-Onlinebefragungen sind Coverage-Probleme (de Leeuw und Hox 2010). Eine Onlinebefragung ist nur dann sinnvoll, wenn alle relevanten Zielgruppen Zugang zu diesem Medium haben und somit eine Übereinstimmung von angestrebter Grundgesamtheit und Befragungspopulation erreicht werden kann (Schnell 2019). Wie bereits in Kapitel 6 aufgeführt, wird das Internet besonders von älteren Personengruppen seltener genutzt. Dies ist problematisch, da sich Patientenbefragungen im Rahmen der gesetzlichen Qualitätssicherung an alle

Altersgruppen richten und ältere Personen entsprechend bei reinen Onlinebefragungen Nachteile bei der Teilnahme hätten und somit wichtige Patientengruppen ausgeschlossen werden würden.

Ein weiterer Nachteil von Single-Mode-Onlinebefragungen ist, dass die Nonresponse-Raten meist höher sind als bei postalischen Befragungen (de Leeuw und Hox 2010, Weigold et al. 2018b, Shin et al. 2012, Börkan 2010, Daikeler 2020, Messer und Dillman 2011, Couper 2011). Dies trifft vor allem dann zu, wenn es sich um Onlinebefragungen mit digitaler Kontaktierung handelt (de Leeuw und Hox 2010, Weigold et al. 2018b, Shin et al. 2012, Börkan 2010, Daikeler 2020, Messer und Dillman 2011, Couper 2011). Die Gründe hierfür sind vielfältig. Da Onlinebefragungen mittlerweile das beliebteste Befragungsinstrument sind, ist es möglich, dass viele Menschen nur schwer zwischen relevanten und nicht relevanten Umfragen unterscheiden können (Daikeler 2020). Zudem werden Onlinebefragungen als unpersönlich und wenig aufwendig für forschende Institutionen wahrgenommen, wodurch sie weniger seriös erscheinen (Daikeler 2020). Ebenfalls können Onlinebefragungen leicht übersehen werden, landen im Spam-Ordner oder werden als weniger verbindlich wahrgenommen als postalische Befragungen (Daikeler 2020, Dillman et al. 2014a). Zudem zeigen Studien, dass Onlinebefragungen größere Zweifel bezüglich des Datenschutzes hervorrufen als postalische Befragungen und es kommt häufiger zu technischen Problemen (Daikeler 2020, Gummer und Daikeler 2020, Dillman 2017).

Eine Möglichkeit zur Verbesserung der Wahrnehmung der Seriosität und damit der Teilnahme rate an Onlinebefragungen bieten bestimmte Designmöglichkeiten, welche in Abschnitt 7.2.1 dargestellt werden. Zudem bieten Mixed-Mode-Befragungen Lösungsansätze, die im folgenden Abschnitt vorgestellt werden.

7.1.2 Mixed-Mode

Mixed-Mode-Befragungen zeichnen sich dadurch aus, dass Fragebogendaten aus verschiedenen Arten der Datenerhebung miteinander kombiniert werden, etwa Daten aus einer Onlinebefragung mit Daten aus einer postalischen Befragung.

Generell ist bei Mixed-Mode-Befragungen zu unterscheiden, ob den Befragten nach dem Erstkontakt die Möglichkeit gegeben wird, die Befragung online oder postalisch zu beantworten (simultan), oder ob sie zuerst nur die Möglichkeit haben, den Fragebogen online zu beantworten und erst bei Nichtbeantwortung nach einer gewissen Zeit einen Papierfragebogen zugeschickt bekommen (sequenziell).²⁴ Die initiale Kontaktierung der Teilnehmerinnen und Teilnehmer erfolgt bei Mixed-Mode-Befragungen üblicherweise postalisch, wobei diese einen Brief mit Link und / oder QR-Code zur Onlinebefragung oder einen Brief mit einem Papierfragebogen erhalten.

²⁴ Im Rahmen von Mixed-Mode-Befragungen lassen sich Onlinebefragungen nicht nur mit postalischen Befragungen kombinieren. So ist es z. B. bei sequenziellen Mixed-Mode-Befragungen üblich, dass man nach der Frist zur Beantwortung des Fragebogens von einer Interviewerin oder einem Interviewer angerufen wird, um die Befragung telefonisch zu beantworten. Da die Patientenbefragungen des IQITG ausschließlich als selbstadministrierte Befragungen durchgeführt werden, wird auf diese Formen nicht weiter eingegangen.

Mixed-Mode – Simultan

Simultane Mixed-Mode-Befragungen zeichnen sich dadurch aus, dass die Befragten nach Erstkontakt die Möglichkeit haben, den Fragebogen über mindestens zwei verschiedene Modi auszufüllen, also online oder postalisch. Hierbei bekommen die Befragten ein postalisches Anschreiben mit einem Fragebogen, in dem zusätzlich ein Link und / oder QR-Code liegt, der die Befragten zur Onlinebefragung führt. Der Vorteil dieses Ansatzes ist, dass die Befragten in ihrem präferierten Modus antworten können. Allerdings ist zu erwarten, dass dieses Vorgehen nicht zu einer ähnlich hohen Kostenreduktion führt wie eine Single-Mode-Onlinebefragung (Vannieuwenhuyze 2014, Stadtmüller und Silber 2021). Gleichzeitig ist nicht zu erwarten, dass die Response-Raten von simultanen postalischen Mixed-Mode-Befragungen mit Web-Option höher sind als von rein postalischen Befragungen (Medway und Fulton 2012, Millar und Dillman 2011). Dies kann u. a. daran liegen, dass der kognitive Aufwand für die Teilnehmerinnen und Teilnehmer höher ist, wenn sie sich für einen Modus entscheiden müssen oder daran, dass der Wechsel von postalischem Anschreiben zu digitaler Fragebogenbeantwortung eine Unterbrechung des Befragungsprozesses darstellt (Medway und Fulton 2012).

Mixed-Mode – Sequenziell (Push-to-Web)

Sequenzielle Mixed-Mode-Befragungen, auch Push-to-Web-Befragungen genannt (Patrick et al. 2020, McMaster et al. 2017, Dillman 2017), zeichnen sich dadurch aus, dass die Befragten nicht von Anfang an die Möglichkeit haben, aus verschiedenen Modi zu wählen. Es wird zuerst versucht, so viele Personen wie möglich über die kostengünstigere Variante zu befragen, d. h. über eine Onlinebefragung (Dillman et al. 2014a). Hat die befragte Person nach einem gewissen Zeitraum den Fragebogen noch nicht online beantwortet, bekommt sie einen Papierfragebogen zugeschickt. Dieser Modus hat den Vorteil, dass er insgesamt kostengünstiger ist als der simultane Mixed-Mode-Ansatz und der klassische postalische Modus (Stadtmüller et al. 2021).

Ob die Teilnehmerinnen und Teilnehmer einer Befragung eine simultane oder sequenzielle Befragung bevorzugen, kann von Faktoren wie etwa vom Alter der Befragten abhängen. So scheinen ältere Menschen eher simultane und jüngere eher sequenzielle Mixed-Mode-Befragungen zu bevorzugen, weshalb es sinnvoll sein kann, eine Kombination aus beidem, sogenanntes Targeting, einzusetzen (Stadtmüller et al. 2021, Callegaro et al. 2015, Freedman et al 2018, Couper 2011). Targeting kann im Rahmen von Mixed-Mode-Befragungen auch genutzt werden, um bestimmte Bevölkerungsgruppen besser zu erreichen, die eine Befragungsart einer anderen vorziehen (siehe Abschnitt 7.1.1).

Vorteile von Mixed-Mode-Befragungen

Prinzipiell können die Vorteile von Single-Mode-Onlinebefragungen, also bspw. die technischen Möglichkeiten wie komplexe Filterführungen oder unmittelbare Datenspeicherung, zu einem gewissen Grad auch auf Mixed-Mode-Befragungen übertragen werden.

Neben den spezifischen, oben dargestellten Vorteilen von simultanen bzw. sequenziellen Mixed-Mode-Befragungen zeigen Studien, dass Mixed-Mode-Befragungen einen Lösungsansatz für Coverage-Probleme darstellen, da durch die unterschiedlichen Befragungsmodi diversere Zielgruppensegmente erreicht werden. So können postalische Befragungen, die zusätzlich zu einer Onlinebefragung angeboten werden, die demographische Repräsentativität eines Samples

erhöhen (DeLeeuw 2018, Cornesse und Bosnjak 2018, Bosnjak et al. 2017). Grundsätzlich ist hierbei jedoch anzumerken, dass es bisher nur wenige empirische Studien zur Effektivität von Mixed-Mode-Befragungen als Lösungsansatz für Coverage-Probleme gibt (DeLeeuw 2018, Couper und Miller 2008).

Wie in Abschnitt 7.1.1 beschrieben, sind die Teilnahmeraten an Single-Mode-Onlinebefragungen meist niedriger als an postalischen Befragungen. Die Studienlage zu der Frage, ob Mixed-Mode-Befragungen dazu geeignet sind, höhere Response-Raten zu erzielen, ist jedoch nicht eindeutig (Millar und Dillman 2011, Schonlau und Couper 2017). So gibt es einerseits Studien, die zeigen, dass Mixed-Mode-Befragungen ähnlich hohe oder sogar höhere Response-Raten als postalische Befragungen erzielen können (Millar und Dillman 2011, Bucks et al. 2020). Dahingegen weisen Schonlau und Couper (2017) in einer Übersichtsarbeit auf die unklare Evidenzlage hin und führen an, dass eine simultan zu einer Paper-Pencil-Befragung angebotene Weboption die Teilnahmerate gegenüber einer rein postalischen Befragung sogar verringern kann. Ebenso ist es laut Schonlau und Couper (2017) unklar, ob sequenzielle Mixed-Mode-Befragungen zu einer Erhöhung der Response-Rate führen.

Nachteile von Mixed-Mode-Befragungen

Das Ziel von Mixed-Mode-Befragungen ist es, Fragebogendaten aus unterschiedlichen Formen der Datenerhebung miteinander zu kombinieren. Grundlage dafür ist die Annahme, dass Daten aus unterschiedlichen Erhebungsarten miteinander vergleichbar sind und Messäquivalenz besteht (de Leeuw und Hox 2010). Ist dies nicht der Fall, können sogenannte „Moduseffekte“ auftreten, d. h. Heterogenität in den Daten, die aufgrund unterschiedlicher Befragungsmodi zustande kommt (Stadtmüller et al. 2021). Moduseffekte können beispielsweise entstehen, wenn Fragen in Onlinebefragungen und postalischen Befragungen unterschiedlich formuliert oder dargestellt werden (de Leeuw und Hox 2010). Studien zeigen jedoch, dass solche Effekte in qualitativ hochwertig designten und durchgeführten Mixed-Mode-Befragungen kaum bedeutsam sind (Stadtmüller et al. 2021, Blom et al. 2015, Börkan 2010) und Messäquivalenz angenommen werden kann (de Leeuw und Hox 2010). Wichtig ist, dass bei der Migration von Papier- zu Onlinefragebögen keine substanziellen Veränderungen bezüglich Frageformat oder -design vorgenommen werden.

7.2 Design von Onlinebefragungen

Grundsätzlich gelten bei der Konstruktion von Onlinebefragungen die gleichen Empfehlungen wie für Paper-Pencil-Befragungen (für einen Überblick siehe Callegaro et al. (2015), Dillman et al. (2014a)). Wie bereits in Abschnitt 7.1.1 aufgeführt, bieten Onlinebefragungen jedoch vielfältigere Designmöglichkeiten. Wichtig ist es, darauf zu achten, dass das Design eines Onlinefragebogens zu den Zielen der Befragung passt und die Vorteile so genutzt werden, dass die Navigation des Fragebogens nutzerfreundlich ist. Im Folgenden werden verschiedene Designmerkmale vorgestellt, die bei der Konstruktion einer Onlinebefragung zu beachten sind.

Design der Startseite

Die Rolle einer Startseite ist es, die Befragten davon zu überzeugen, an der Onlinebefragung teilzunehmen. Inhalt und Gestaltung der Startseite sind daher von großer Bedeutung. Die Startseite sollte ansprechend und übersichtlich gestaltet sein und die wichtigsten Informationen zur Umfrage wie Institution, Zweck, Inhalt, Datenschutz sowie Kontaktmöglichkeiten beinhalten (Callegaro et al. 2015, Haer und Meidert 2013).

Scrolling vs. Paging

Bei Onlinebefragungen können die einzelnen Items auf unterschiedliche Weise über die Seiten des Fragebogens verteilt werden. Es gibt zum einen die Möglichkeit eines „Scrolling“-Designs (Callegaro et al. 2015, Vincente und Reis 2010), bei dem die Fragen auf einer einzelnen Webseite angezeigt werden und die Befragten von Frage zu Frage scrollen, um die Items zu beantworten. Am Ende der Seite befindet sich eine Schaltfläche, mithilfe derer die Antworten übermittelt werden können. Zum anderen gibt es die Möglichkeit eines „Paging“- bzw. „Screen“-Designs (Callegaro et al. 2015, Vincente und Reis 2010), bei dem die Items auf mehrere Seiten verteilt werden und sich auf jeder einzelnen Seite eine Schaltfläche zum Übermitteln der Antworten befindet. Hierbei werden die Antworten gespeichert, bevor die nächste Frage angezeigt wird. Studien zeigen, dass die Bearbeitungszeiten von Befragungen im Scrolling-Design meist kürzer sind (Callegaro et al. 2015, De Bruijne und Wijnant 2014, Mavletova und Couper 2016). Jedoch besteht hier die Gefahr höherer Item-Nonresponse-Raten im Vergleich zu Paging-Designs (Vincente und Reis 2010, Mavletova und Couper 2016) und es kann leichter passieren, dass Daten verloren gehen, da die Itemantworten nicht wie im Paging-Design direkt nach der Bearbeitung einer Frage gespeichert werden, sondern erst gesammelt am Ende der Umfrage (Callegaro et al. 2015). Um diese Nachteile zu vermeiden, wird in der Literatur das Paging-Design empfohlen (Callegaro et al. 2015).

Fortschrittsanzeigen

Ein weiterer zu berücksichtigender Punkt ist die Einbindung von Fortschrittsanzeigen in Onlinebefragungen. Bei der Bearbeitung eines Fragebogens in Papierform können die Teilnehmerinnen und Teilnehmer ihren Fortschritt einfach abschätzen. Bei Onlinebefragungen können Fortschrittsanzeigen in grafischer Form als Balken oder in numerischer Form als Prozentangabe eingebunden werden, die den Befragten zeigen, zu welchem Teil sie die Befragung schon abgeschlossen haben. Diese Anzeigen können entweder permanent eingeblendet werden oder intermittierend zwischen verschiedenen Umfrageteilen. Außerdem können Fortschrittsanzeigen unterteilt werden in konstante bzw. lineare Anzeigen, „Fast-to-slow“- und „Slow-to-fast“-Anzeigen (Villar et al. 2013). Bei konstanten Fortschrittsanzeigen steigt der angezeigte Fortschritt linear zu den bearbeiteten Items im Fragebogen. Bei Fast-to-slow-Anzeigen verlangsamt sich die Geschwindigkeit des Fortschritts im Laufe der Fragebogenbeantwortung, wobei der Fortschritt auf den ersten Bildschirmseiten schnell ist und sich dann verlangsamt. Bei Slow-to-fast-Anzeigen hingegen beschleunigt die Fortschrittsgeschwindigkeit zum Ende hin. Ziel von Fortschrittsanzeigen ist es, vorzeitige Abbrüche der Befragung zu vermeiden. Die Evidenz zum Effekt solcher Anzeigen auf Befragungsabbrüche ist jedoch nicht eindeutig (Liu und Wronski 2018) und mögliche Effekte sind von mehreren Faktoren abhängig. So beeinflussen beispielsweise die Länge des Fra-

gebogens und die Erwartungen der Teilnehmerinnen und Teilnehmer hinsichtlich der Ausfülldauer, ob ein Fortschrittsbalken vorzeitige Abbrüche verhindert oder nicht (Yan et al. 2011, Vincente und Reis 2010). Studien zeigen zudem, dass Dauer und Zeitpunkt des Einblendens einer Fortschrittsanzeige sowie die Art der Fortschrittsanzeige (konstant bzw. Fast-to-slow oder Slow-to-fast) beeinflussen können, ob die Umfrage vorzeitig abgebrochen wird oder nicht (Vincente und Reis 2010, Villar et al. 2013). Das Einbinden von Fortschrittsanzeigen in Onlinebefragungen ist entsprechend von mehreren Faktoren abhängig und ist für den spezifischen Fall anhand der vorliegenden Erkenntnisse zu prüfen.

Eingabeaufforderung

Bei Onlinebefragungen besteht im Gegensatz zu Paper-Pencil-Befragungen die Möglichkeit, sogenannte Eingabeaufforderungen (engl. „Nonresponse-Prompts“) einzubauen, um zu vermeiden, dass die Teilnehmerinnen und Teilnehmer einzelne Fragen unbeantwortet lassen (Couper 2008). Bei der Gestaltung dieser Eingabeaufforderung sollte darauf geachtet werden, dass sie freundlich und kurz formuliert ist und es sollte abgewogen werden, ob es sich lediglich um eine einmalige Aufforderung zur Eingabe handeln soll (sog. „soft Nonresponse-Prompts“), oder ob die weitere Bearbeitung der Befragung bei Nichtbeantwortung eines Items nicht mehr möglich sein soll (sog. „hard Nonresponse-Prompts“) (Callegaro et al. 2015). Die Einbindung letzterer Variante sollte gründlich abgewogen werden. Die Vorteile bestehen darin, dass Item-Nonresponse vermieden wird und die Datenauswertung vereinfacht wird, da keine Analyse zu fehlenden Daten durchgeführt werden muss (Callegaro et al. 2015). Jedoch haben „hard Nonresponse-Prompts“ den Nachteil, dass sie als „Zwang“ zur Beantwortung wahrgenommen werden und dadurch zu höheren Abbruchraten führen könnten (Callegaro et al. 2015).

Item- und Antwortformate

Prinzipiell bieten Onlinebefragungen zahlreiche Möglichkeiten, Items darzustellen und sind damit flexibler gestaltbar als Papierfragebögen. So können die Fragen einzeln nacheinander dargestellt werden, mehrere Items können als Raster kombiniert oder als Filterfrage dargestellt werden. Bei Filterfragen werden die Teilnehmerinnen und Teilnehmer abhängig von ihrer Antwort zu einer Folgefrage weitergeleitet. Im Rahmen von Items im Rasterformat werden mehrere Fragen mit gleichem Antwortformat zusammengefasst und in Zeilen oder Spalten dargestellt. Studien zeigen, dass das Format von Items und Antwortoptionen einen Effekt auf Bearbeitungszeit, Item-Nonresponserate, Abbruchrate oder den Messfehler von Umfragen haben kann (Debell et al. 2021, Mavletova et al. 2017, Liu und Cernat 2016). Auch die Position, in der Items auf einem Screen dargestellt werden oder die horizontale bzw. vertikale Ausrichtung der Antwortoptionen, kann einen Einfluss auf das Antwortverhalten der Befragten haben (Tourangeau et al. 2013, Hu 2020, Lenzner et al. 2014, Revilla und Couper 2018).

Auch zur Antwortauswahl in Onlinebefragungen können unterschiedliche Formate gewählt werden. Hierzu gehören bspw. Radio-Buttons und Drop-Down-Menüs, die sich für eine Einfachauswahl eignen, oder freie Texteingabefelder. Möglich sind ferner Skalen, Matrizen, Mehrfachauswahlen, Reihenfolgen und vielfältige vorgefertigte Fragetypen. Bei der Nutzung von Radio-Buttons werden alle Antwortmöglichkeiten auf einem Screen angezeigt und die Teilnehmerinnen und Teilnehmer wählen eine Antwortoption. Bei Drop-Down-Menüs wird die Liste der Antwortmöglichkeiten erst nach einem Klick auf das Menü angezeigt und die Antwort kann dann

ausgewählt werden. Bei freien Texteingabefeldern können die Befragten ihre Antwort selbst eintippen. Dieses Format eignet sich besonders für offene Fragen. Da diese im Rahmen von Patientenbefragungen der gesetzlichen Qualitätssicherung nicht verwendet werden, wird dieses Antwortformat hier nicht empfohlen. Eine Übersichtsarbeit von Vicente und Reis (2010) zeigt, dass die Verwendung von Radio-Buttons anstelle von Drop-Down-Menüs oder freier Texteingabe bei geschlossenen Fragen zu niedrigeren Item-Nonresponse-Raten führt, das Antwortformat jedoch keinen Effekt auf die generelle Abbruchrate von Umfragen zu haben scheint (Vicente und Reis 2010).

Device-Effekte

Onlinebefragungen können mithilfe diverser internetfähiger Endgeräte bearbeitet werden, bspw. PCs, Smartphones oder Tablets. Hier ist zu beachten, dass das Design der Befragung für das jeweilige Endgerät optimiert wird und der Fragebogen auf allen Endgeräten ähnlich aussieht, um die Vergleichbarkeit der Daten sicherzustellen und sogenannte „Device-Effekte“ zu vermeiden (Antoun et al. 2017, Antoun et al. 2018, Dillman 2017).

Im Rahmen einer systematischen Literaturrecherche evaluieren Antoun et al. (2018) den Effekt unterschiedlicher Designmerkmale von Smartphonebefragungen auf die Antwortqualität von Umfragen. Laut den Autoren soll insbesondere die Optimierung von Onlinebefragungen für Smartphones berücksichtigt werden, um Messunterschiede zu vermeiden. Der Fragetext solle ausreichend groß sein, um einfach gelesen werden zu können und der Button zur Antwortauswahl groß genug sein, um eine einfache Antworteingabe zu ermöglichen. Zudem solle die Größe der Fragen an die Breite von Smartphonebildschirmen angepasst werden, damit die Frage auf einer Bildschirmseite angezeigt und somit horizontales Scrollen vermieden werden kann. Die Autoren empfehlen ebenfalls die Nutzung einfacher Fragetypen wie Single- oder Multiple-Choice-Fragen gegenüber komplizierten Fragetypen im Raster-, Dropdown- oder Slider-Format und weisen darauf hin, dass die Befragung auf diversen Endgeräten gleich zu bedienen sein sollte. Andere Studienergebnisse legen zudem nahe, komplexe grafische Designs zu vermeiden, um keine technischen Probleme zu verursachen und die Teilnehmerinnen und Teilnehmer nicht abzulenken (Vicente und Reis 2010).

Zahlreiche Studien zeigen, dass die Datenqualität von Umfragen, die auf verschiedenen Endgeräten durchgeführt werden, vergleichbar hoch und Messäquivalenz gegeben ist, solange die Befragung für das jeweilige Endgerät optimiert und nutzerfreundlich gestaltet ist (Antoun et al. 2017, Antoun et al. 2019, Mavletova und Couper 2015, Tourangeau et al. 2017b, Tourangeau et al. 2017a, Schlosser und Mays 2017, Sommer et al. 2016, Wells et al. 2013, Toninelli und Revilla 2016, Lugtig und Toepoel 2016, Keusch und Yan 2017).

7.2.1 Besonderheiten beim Design von Mixed-Mode-Befragungen

Prinzipiell ist beim Design von Mixed-Mode-Befragungen zu beachten, dass sich die Formate und Formulierung der Fragen und Antworten zwischen Online- und Papierversion nicht unterscheiden (Stadtmüller et al. 2021, Dillman et al. 2014a). Zudem können Moduseffekte entstehen, wenn sich Papier- und Onlinefragebogen in der grafischen Darstellung zu stark unterscheiden, d. h. es sollte das gleiche Design verwendet werden (Dillman et al. 2014a).

Ferner sollte, wie in Abschnitt 7.1.2 bereits erwähnt, beachtet werden, dass die technischen Möglichkeiten von Onlinebefragungen nur so umgesetzt werden, dass sie den Teilnehmerinnen und Teilnehmern die Navigation des Fragebogens erleichtern oder Fehler reduzieren, jedoch nicht, wenn sie möglicherweise zu Messunterschieden führen. Ein Beispiel hierfür wäre etwa die Nutzung von Filterführungen oder Plausibilitätsprüfungen bei der Onlineversion der Befragung, die den Befragten die Navigation erleichtern und Fehleranfälligkeit reduzieren. Die Einbindung von Multimediainhalten in die Onlinebefragung hingegen ist im Rahmen einer Mixed-Mode-Befragung weniger empfehlenswert, da dies in Papierfragebögen nicht in gleichem Rahmen möglich ist und sich somit Unterschiede in den Modi ergeben würden.

Besonders wichtig ist es bei Mixed-Mode-Befragungen, dass im Entwicklungsprozess explizit geprüft wird, ob die kombinierten Erhebungsarten vergleichbare Daten liefern. Ist dies sichergestellt, ist die Gefahr von Moduseffekten als gering anzusehen (siehe Abschnitt 7.1.2).

7.3 Bewertung der Umsetzung von onlinebasierten Patientenbefragungen im Regelbetrieb

7.3.1 Datenfluss Patientenbefragung gemäß DeQS-RL

Die Vorgaben für die Umsetzung einer Patientenbefragung werden in § 299 SGB V definiert und durch die Regelungen der DeQS-RL konkretisiert. Es handelt sich hierbei um eine Single-Mode-Befragung unter Nutzung von papierbasierten Fragebögen. Zur Bewertung der Umsetzbarkeit der in Abschnitt 7.1 skizzierten Möglichkeiten einer onlinebasierten Patientenbefragung, insbesondere im Hinblick auf die Veränderung des Datenflusses, organisatorischer Abläufe und Aufgaben der Versendestelle Patientenbefragung (VPB) ist ein Verständnis der aktuellen Vorgaben notwendig (siehe Abbildung 14).

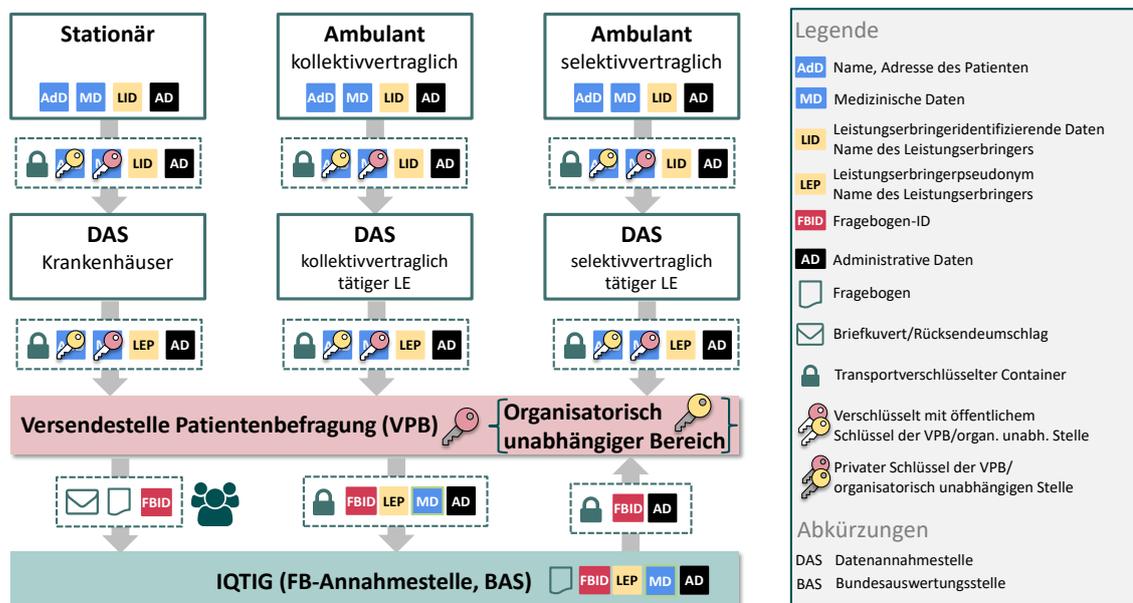


Abbildung 14: Datenfluss Patientenbefragung gemäß DeQS-RL

Der **Leistungserbringer** (LE) unterliegt einer Dokumentationspflicht der sektorenübergreifenden Qualitätssicherung nach § 136 ff. SGB V, in diesem Fall ist er insbesondere zu einer Beteiligung an Patientenbefragungen verpflichtet. Er stellt alle für die korrekte Auswahl und den Versand des Fragebogens notwendigen Informationen zusammen und übermittelt diese an die VPB. Der Leistungserbringer exportiert die Adressdaten der Patientin oder des Patienten (AdD), die medizinischen Daten zur Auswahl des korrekten Fragebogens in der Versendestelle Patientenbefragung (MD), die leistungserbringeridentifizierenden Daten (LID) sowie administrative Daten (AD). Die Adressdaten und medizinischen Daten werden mit zwei unterschiedlichen kryptografischen Schlüsseln verschlüsselt, die jedoch beide zur Versendestelle Patientenbefragung gehören. Anschließend werden die Daten an die Datenannahmestelle auf Landesebene (DAS) übermittelt.

Die Aufgabe der **Datenannahmestelle** ist die Pseudonymisierung der leistungserbringeridentifizierenden Daten. Sie hat keine Einsicht in die Adressdaten oder die medizinischen Daten. Sie übermittelt die Daten mit dem Leistungserbringerpseudonym weiter an die Versendestelle Patientenbefragung.

Die **Versendestelle** Patientenbefragung nimmt die Daten der Datenannahmestellen entgegen und entschlüsselt diese gemäß den Vorgaben der DeQS-RL. Die medizinischen Daten werden sofort bei der Datenannahme entschlüsselt und geprüft. Die Adressdaten werden vorerst verschlüsselt in einem separaten Bereich der Versendestelle Patientenbefragung vorgehalten und nur bei Bedarf nach erfolgter Stichprobenziehung entschlüsselt. Die Versendestelle Patientenbefragung identifiziert die Patientendatensätze, die zur Teilnahme an der Befragung bestimmt wurden und wählt ggf. die korrekte Fragebogenversion aus. Anschließend erfolgt die Entschlüsselung der zugehörigen Adressinformationen und die Fertigstellung des Fragebogens und des Anschreibens. Jedem Fragebogen wird eine zufällige, eindeutige Identifikationsnummer zugeordnet. Die Fragebogen-ID wird für eine Rücklaufkontrolle sowie die Unterdrückung von Erinnerungsschreiben und die Übermittlung von Daten von der Versendestelle an das IQTIG genutzt. Nach der Stichprobenziehung, der Fragebogenauswahl und dem Druck des Fragebogens und des Anschreibens und der Kuvertierung werden die Fragebögen an die Patientinnen und Patienten postalisch versendet.

Die durch die Stichprobenziehung ausgewählte **Patientin bzw. der Patient** erhält von der Versendestelle per Post einen Fragebogen und ggf. ein Erinnerungsschreiben. Er bzw. sie füllt den Fragebogen aus und sendet ihn mit dem beiliegenden Rückumschlag an die Fragebogenannahmestelle des IQTIG zurück.

Die **Fragebogenannahmestelle** des IQTIG nimmt die Fragebögen entgegen, liest diese ein und übermittelt die Datensätze in abgestimmter Regelmäßigkeit an die Bundesauswertungsstelle. Ihr obliegt außerdem die datenschutzkonforme Lagerung und Vernichtung der Fragebögen. Die Fragebogenannahmestelle übermittelt die Fragebogen-IDs der eingegangenen Fragebögen zur Steuerung der kontrollierten Erinnerungsschreiben an die Versendestelle.

7.3.2 Kontaktaufnahme bei onlinebasierten Befragungen

Die Art der Kontaktaufnahme (postalisch/elektronisch) spielt eine entscheidende Rolle bei der Beurteilung der Umsetzbarkeit der verschiedenen Befragungsmodelle. Die verschiedenen Vor- und Nachteile sowie die demzufolge notwendigen Änderungen der Prozesse der Versendestelle Patientenbefragung sind im Folgenden dargestellt. Betrachtet werden in diesem Zusammenhang nur die Möglichkeiten, die für die gesetzliche verpflichtende Qualitätssicherung grundsätzlich in Frage kommen (siehe hierzu auch Abschnitt 7.1).

7.3.2.1 Postalische Kontaktaufnahme

Die postalische Kontaktaufnahme zur Patientin oder zum Patienten erfolgt auf gleichem Wege wie dies bei der postalischen Befragung gemäß DeQS-RL etabliert ist. In einem Anschreiben erhält die Patientin oder der Patient die Zugangsdaten zur Onlinebefragung.

Vorteile bei einer postalischen Kontaktaufnahme

- Die Nutzung der Adressdaten zur Kontaktaufnahme der Patientin oder des Patienten sind bereits in den gesetzlichen Grundlagen verankert und datenschutzrechtlich geprüft.
- Die Adressdaten können automatisiert von der eGK-Versichertenkarte der Patientin oder des Patienten ausgelesen und genutzt werden, das Fehlerrisiko wird minimiert.
- Eine zusätzliche Erhebung von patientenidentifizierenden Informationen wie z. B. E-Mail-Adresse oder Telefonnummer wird vermieden.
- Das Fehlerrisiko bei einer manuellen Erhebung der zusätzlichen Informationen durch das medizinische Personal wird minimiert und eine erhöhte Arbeitsbelastung vermieden.
- Einladungen zur Teilnahme an der Befragung auf elektronischem Wege können verloren gehen, z. B. durch SPAM-Filter o. ä., für Briefe ist dies eher unwahrscheinlich.

Herausforderungen bei einer postalischen Kontaktaufnahme

- Die postalische Kontaktaufnahme stellt bei einer Onlinebefragung einen Medienbruch dar, d. h. die Patientin oder der Patient muss die Informationen über die Internetadresse der Onlinebefragung (URL) sowie die Zugangsdaten von der Papierform in eine elektronische Form übertragen. Hierfür gibt es u. a. folgende Möglichkeiten:
 - Nutzung eines QR-Codes: Alle Angaben (URL inkl. Zugangsdaten) können in einem QR-Code gespeichert werden. Die Patientin bzw. der Patient benötigt ein Smartphone oder Tablet sowie eine entsprechende Software, um den QR-Code einlesen zu können. Die URL wird dann direkt auf dem Smartphone oder Tablet im Webbrowser geöffnet.
 - Nutzung einer URL sowie separate Zugangsdaten: Die Patientin bzw. der Patient muss die URL manuell auf ihrem PC, Smartphone oder Tablet abtippen um die Internetseite der Befragung zu öffnen. Anschließend müssen die Zugangsdaten manuell abgetippt werden.
 - Nutzung einer URL inkl. Zugangsdaten: Diese URL ist durch die integrierten Zugangsdaten sehr lang und kryptisch. Abhilfe kann an dieser Stelle ein Dienst zur Kürzung der URLs schaffen (siehe Abschnitt 7.3.5.4).
- Der Versand von Briefen ist kostenintensiv.

- Aus ökologischer Sicht ist der Briefversand ressourcenaufwändiger.

Prozesse in der Versendestelle Patientenbefragung

Bei einer postalischen Kontaktaufnahme können die bisher durch den G-BA etablierten Prozesse weitestgehend genutzt werden. Die Aufgaben und Prozesse der Versendestelle Patientenbefragung würden sich nach aktuellem Kenntnisstand folgendermaßen gestalten:

- Die Versendestelle Patientenbefragung führt die Stichprobenziehung durch.
- Die Versendestelle Patientenbefragung erstellt die Zugangsdaten für das Befragungsportal oder erhält diese vom Anbieter des Befragungsportals. Auf Basis der medizinischen Daten wird der Zugang zum entsprechenden QS-Verfahren und Fragebogentyp erstellt.
- Auf dem Anschreiben bzw. auf einem separaten Schreiben werden die Zugangsdaten für die Patientin oder den Patienten zur Onlinebefragung aufgedruckt.
- Im Falle einer Onlinebefragung im Single-Mode: Die Auswahl und der Druck der Fragebögen würde obsolet werden, es würden lediglich die Anschreiben gedruckt und versendet werden.

7.3.2.2 Elektronische Kontaktaufnahme

Die elektronische Kontaktaufnahme mit der Patientin oder dem Patienten kann per E-Mail, SMS, Messenger oder Smartphone-App erfolgen.

Vorteile einer elektronischen Kontaktaufnahme

- Es entstehen keine Medienbrüche; die Patientin bzw. der Patient erhält eine URL und kann diese anklicken um den Fragebogen direkt zu beantworten.
- Die Nutzung elektronischer Wege ist kostensparend.
- Die Vermeidung des Briefversands ist ökologisch nachhaltig.

Nachteile einer elektronischen Kontaktaufnahme

- Nicht alle Patientinnen und Patienten, insbesondere der höheren Altersstufen, besitzen einen Internetzugang, ein Smartphone oder eine E-Mail-Adresse (siehe Kapitel 6)
- Die Einladungen können verloren gehen, indem sie z. B. durch das E-Mail-Programm als Spam erkannt und in einen Spam-Ordner verschoben werden.
- Im Rahmen der Befragung müssten weitere patientenidentifizierende Daten erhoben werden, die so aktuell im Klinikalltag nicht standardmäßig erhoben werden bzw. auf der eGK-Versichertenkarte gespeichert sind.

Prozesse in der Versendestelle Patientenbefragung

Bei einer ausschließlich elektronischen Kontaktaufnahme entstehen weitreichende Änderungen im Aufgabenspektrum der Versendestelle im Vergleich zu einer papierbasierten Kontaktaufnahme:

- Das Element „Adressdaten“, welches vom Leistungserbringer an die Versendestelle Patientenbefragung geschickt wird, enthält den Namen und jeweils die E-Mail-Adresse oder Telefonnummer, abhängig vom gewünschten Kontaktweg. Die Angaben zu Straße und Ort können entfallen.
- Die Versendestelle Patientenbefragung führt die Stichprobenziehung durch.
- Die Versendestelle erstellt die Zugangsdaten für das Befragungsportal oder erhält diese vom Anbieter des Befragungsportals. Auf Basis der medizinischen Daten wird der Zugang zum entsprechenden QS-Verfahren und Fragebogentyp erstellt.
- Für alle in die Stichprobe eingeschlossenen Patientinnen und Patienten werden die „Adressdaten“, also Name und E-Mail-Adresse, entschlüsselt. Anschließend wird eine E-Mail mit einem Anschreibetext und den Zugangsdaten verschickt.
- Es wäre ausschließlich eine Onlinebefragung im Single-Mode möglich. Die Auswahl und der Druck der Fragebögen und Anschreiben würde daher obsolet werden.

7.3.3 Umsetzbarkeit von onlinebasierten Single-Mode-Befragungen

Die Umsetzung von onlinebasierten Single-Mode-Befragungen erfordert Anpassungen bei den Prozessen der Versendestelle Patientenbefragung als auch bei der Fragebogenannahmestelle. Die Versendestelle Patientenbefragung ist in diesem Szenario nicht mehr für den Druck und den Versand des Fragebogens zuständig. Die weiteren durch die Versendestelle Patientenbefragung gemäß DeQS-RL definierten Aufgaben (z. B. Stichprobenziehung und Druck des Anschreibens) bleiben davon unberührt. Zusätzlich zu den bisherigen Aufgaben sind durch die Versendestelle die Zugangsdaten zum Webportal auf das Anschreiben aufzudrucken.

Die Prozesse in der Fragebogenannahmestelle ändern sich bei der Durchführung einer onlinebasierten Single-Mode-Befragung erheblich. Die Aufgaben der Fragebogenannahmestelle verändern sich von einem Dienstleistungsanbieter zur Digitalisierung von Papierfragebögen hin zu einem IT-Serviceprovider für die Durchführung webbasierter Befragungen. Das Aufgabenspektrum der beiden Stellen ist nahezu komplett unterschiedlich. Die hierfür notwendigen Anforderungen sind detailliert in Abschnitt 7.3.5 dargestellt.

Die Umsetzung von onlinebasierten Single-Mode-Befragungen erscheint aktuell grundsätzlich möglich. Aufgrund der geringen Änderungen der Prozesse der Versendestelle Patientenbefragung können die entsprechenden Anforderungen ggf. über eine Anpassung des Auftrags mit dem Dienstleister der Versendestelle Patientenbefragung realisiert werden. Eine Änderung des Aufgabenspektrums der vom IQTIG vorgesehenen Fragebogenannahmestelle ist im Hinblick auf eine onlinebasierte Patientenbefragung jedoch aus Sicht des IQTIG voraussichtlich nicht ohne weiteres möglich. Für diese Aufgabe bedarf es ggf. der Etablierung einer eigenen Annahmestelle für onlinebasierte Patientenbefragungen.

7.3.4 Umsetzbarkeit von Mixed-Mode-Befragungen

In Abschnitt 7.1.2 werden verschiedene Möglichkeiten zur Umsetzung von Mixed-Mode-Befragungen im Hinblick auf die Bereitstellungszeitpunkte des Papierfragebogens und des Zugangs zum Onlinefragebogen skizziert. Dies führt aus Umsetzungs- und Datenflusssicht zu einer Vielfalt

an Lösungsmöglichkeiten und -alternativen, deren detaillierte Ausarbeitung nicht im Rahmen dieses Berichts erfolgen kann. Es werden die allgemeinen Zusammenhänge und Herausforderungen bei der Durchführung einer Mixed-Mode-Befragung erläutert.

Die Prozesse und Aufgaben in der Versendestelle Patientenbefragung erfahren durch eine Mixed-Mode-Befragung geringfügige Änderungen. Diese Änderungen gestalten sich jedoch komplexer als bei der Single-Mode-Befragung. Neben dem Druck des Anschreibens mit den Zugangsdaten zur Onlinebefragung muss der Fragebogen in Papierform zusätzlich bereitgestellt und versendet werden. Die Vorgaben, zu welchen Zeitpunkten ausschließlich die Zugangsdaten zur Onlinebefragung, ausschließlich der Fragebogen in Papierform oder eine Kombination aus beiden Befragungsinstrumenten an die Patientin oder den Patienten versandt werden sollen, verursachen in der Versendestelle Patientenbefragung einen erhöhten Aufwand.

Die in Abschnitt 7.3.3 ausgeführten Informationen zu einer Fragebogenannahmestelle für Onlinebefragungen gelten ebenso für Mixed-Mode-Befragungen. Bei Mixed-Mode-Befragungen besteht für die Patientin oder den Patienten die Möglichkeit, sowohl einen Papierfragebogen auszufüllen und einzusenden als auch die Möglichkeit der Beantwortung des Onlinefragebogens. Abhängig von der Konzeption der Mixed-Mode-Befragung besteht ggf. die Möglichkeit, dass die Patientin oder der Patient beide Befragungsinstrumente ausfüllt. Diesem Umstand muss bei der Konzeption der Befragung Rechnung getragen werden. Dazu sind unter anderem folgende Fragen zu beantworten:

- Welcher Fragebogen wird in die Auswertung eingeschlossen und welcher nicht?
- Wird eine Kommunikation zwischen Fragebogenannahmestelle für Papierfragebögen und Fragebogenannahmestelle Onlinebefragungen etabliert, sodass bspw. nach Eingang eines Papierfragebogens der Zugang zum Webportal deaktiviert wird?

Die Umsetzung von Mixed-Mode-Befragungen erscheint möglich. Auch für deren Umsetzung ist eine Erweiterung des Leistungsspektrums der Versendestelle Patientenbefragung sowie die Einrichtung einer Annahmestelle für Onlinebefragungen notwendig, wie dies bereits in Abschnitt 7.3.3 ausgeführt wurde.

7.3.5 IT-technische Umsetzung von Onlinebefragungen mittels eines Webportals

An ein künftiges Webportal für die Bereitstellung von Onlinebefragungen ergehen folgende funktionale und nicht funktionale Anforderungen.

7.3.5.1 Funktionale Anforderungen an das Webportal

Tabelle 5: Funktionale Anforderungen an ein Webportal für Onlinebefragungen

Lfd. Nr.	Anforderung	Erläuterung
F-001	Multiuserfähigkeit	Das Webportal lässt es zu, dass sich mehrere Umfrageadministratoren gleichzeitig am System anmelden können und Umfragen bearbeiten können.

Lfd. Nr.	Anforderung	Erläuterung
F-002	Benutzerverwaltung	Die Benutzerinnen und Benutzer können über eine integrierte Benutzerverwaltung angelegt werden.
F-003	Gleichzeitige Bereitstellung mehrerer Umfragen	Das Webportal erlaubt es, dass mehrere Befragungen gleichzeitig aktiv sind.
F-004	Bereitstellung von Fragetypen	Das Webportal stellt eine Vielzahl von vorgefertigten Fragetypen bereit: <ul style="list-style-type: none"> ▪ Textfragen ▪ Einfachauswahl ▪ Mehrfachauswahl ▪ Matrizen ▪ Skalen ▪ Sortierungen
F-005	Konditionale Verknüpfung von Fragen	Das Webportal unterstützt das Anzeigen von Fragen in Abhängigkeit von vorherigen Eingaben (Konditionen, Sprungmarken).
F-006	Plausibilitätsprüfungen	Das Webportal erlaubt weiche und harte Plausibilitätsprüfungen: <ul style="list-style-type: none"> ▪ Weich: es wird eine Warnung ausgegeben, aber die Eingabe wird toleriert ▪ Hart: ein Verstoß gegen die Vorgaben zur Eingabe bspw. Wertebereiche führt dazu, dass die Befragung nicht fortgesetzt werden kann, ehe der Fehler korrigiert wurde
F-007	Fehlermeldungen	Die Hinweise und Fehlermeldungen sind verständlich und hilfreich oder können redaktionell editiert werden
F-008	Layoutoptionen	Das Webportal bietet die Möglichkeit, das Layout der Weboberfläche der Befragung anzupassen (Customizing, Templating): <ul style="list-style-type: none"> ▪ Individuelle Layouts des Kopfbereiches (Header) und der gesamten Umfrage, sodass das Design dem der Papierbögen angepasst werden kann ▪ Unterstützung von Scrolling und Paging ▪ Platzierbarkeit einer Fortschrittsanzeige ▪ Platzierbarkeit von Hilfstexten ▪ Konfigurierbare Druckansicht
F-009	Unterstützung von anonymen Befragungen	Es können anonyme Befragungen angelegt werden, die von jeder Patientin / jedem Patienten in Besitz der URL einmal pro Endgerät oder IP-Adresse ausgefüllt werden können.

Lfd. Nr.	Anforderung	Erläuterung
F-010	Unterstützung von individuellen Umfragen	Es können geschützte Befragungen angelegt werden, deren Teilnehmerinnen und Teilnehmer ... <ul style="list-style-type: none"> ▪ ... mittels eines Austauschformates importiert werden können ▪ ... sich für die Befragung registrieren können ▪ ... sich mit einer Benutzername/Passwort-Kombination oder einem Zugangsschlüssel identifizieren können ▪ ... über eine individuelle URL auf die Befragung zugreifen ▪ ... die für nicht anonyme Befragungen eine Antwortpersistierung bietet und mit dieser das Zwischenspeichern der Befragung erlaubt.
F-011	Testbarkeit	Die Befragungen müssen getestet werden können, bevor sie zu einem gewissen Zeitpunkt aktiviert werden.
F-012	Exportierbarkeit der Umfragedaten	Die Daten der ausgefüllten Fragebögen müssen mit einer Zuordnung zum Benutzer so gespeichert werden, dass Sie im Anschluss mit einer Statistiksoftware ausgewertet werden können.

Alle Anforderungen sind durch am Markt erhältliche vorgefertigte Befragungslösungen grundsätzlich realisierbar, wobei aus dem Aspekt der individuellen Zugangslösung der größte Konzeptions- und Umsetzungsaufwand resultiert.

7.3.5.2 Nichtfunktionale Anforderungen an die Software

Neben den funktionalen Anforderungen muss das Webportal auch die folgenden nichtfunktionalen Anforderungen zu leisten haben:

Tabelle 6: Nicht funktionale Anforderungen an ein Webportal für Onlinebefragungen

Lfd. Nr.	Anforderung	Erläuterung
N-001	Security	Das Webportal muss aktuellen und standardgemäßen Sicherheitsanforderungen genügen um zu vermeiden, dass ein unbefugter Zugriff auf die Fragebögen oder die Befragungsdaten stattfinden kann.
N-002	Performance	Das Webportal muss es erlauben bzw. mittels Clustering und Loadbalancing so erweiterbar sein, dass bis zu 50.000 Benutzerinnen und Benutzer gleichzeitig ihre Fragebögen bearbeiten können.

Lfd. Nr.	Anforderung	Erläuterung
		Das Webportal muss in der Lage sein, Einladungen zeitlich gestaffelt (bspw. 1.000 pro Stunde) zu versenden, um die Belastung von E-Mail-Systemen kontrollierbar zu halten.
N-003	Kompatibilität	Das System muss weitgehend unabhängig von Hardware und Software funktionieren; das gilt insbesondere für verschiedene Eingabegeräte wie Laptop, Tablet oder Smartphone.
N-004	Usability	Bereitstellung eines URL-Kürzungsdienstes. Der Dienst wandelt lange URLs in kurze Zeichenfolgen um, die vom Befragten unkompliziert von einem papierbasierten Anschreiben in den Webbrowser übertragen werden kann. Die gekürzte URL lenkt den Befragten auf die Webseite der Patientenbefragung und enthält ggf. schon die Zugangsdaten für eine Anmeldung am System.

Auch diese Anforderungen sind bei einer entsprechenden Konzeption und Skalierung erfüllbar.

7.3.5.3 Projektanforderungen

Schließlich bestehen an das Webportal Anforderungen an die Projektierbarkeit.

Tabelle 7: Projektanforderungen an ein Webportal für Onlinebefragungen

Lfd. Nr.	Anforderung	Erläuterung
P-001	Zeitraum	Die Umsetzung und Publizierung einer Umfrage ist binnen 4 Wochen möglich.

7.3.5.4 Nutzung eines URL-Kürzungsdienstes

Dienste zur Kürzung von Internetadressen existieren seit vielen Jahren und werden durch kommerzielle Anbieter, meist im Rahmen ihrer eigenen Plattformen und sozialen Netzwerke, angeboten. Sie ermöglichen es eine lange Internetadresse, bspw.

<https://iqtig.org/qs-berichte/patientenbefragung-schizophrenie/>

in eine kurze Adresse zu kürzen. Der URL-Shortener von Twitter erstellt daraus folgende URL:

<https://t.co/9egZF8Uvff>

Diese kurzen URLs können von Patientinnen und Patienten im Rahmen der Onlinebefragung einfacher im Webbrowser eingegeben werden. Insbesondere bei komplizierten, langen URLs, wie sie bei einer inkludierten Übermittlung der Zugangsdaten entstehen, ist die Nutzung eines Kürzungsdienstes sinnvoll.

Öffentlich verfügbare Kürzungsdienste sollten an dieser Stelle nicht genutzt werden, da diese den Nutzer tracken und ggf. sensible Daten erfassen können. Ein entsprechender Dienst könnte

z. B. durch den Betreiber des Onlineportals bereitgestellt werden. Es existieren verschiedene Open-Source-Projekte, die kostenfrei genutzt und im eigenen Rechenzentrum betrieben werden können. Alternativ dazu kann ein entsprechender URL-Shortener eigenentwickelt werden.

7.4 Darstellung der gesetzlichen und datenschutzrechtlichen Rahmenbedingungen

Die rechtlichen Rahmenbedingungen zur Durchführung einer Onlinebefragung in der gesetzlichen Qualitätssicherung ergeben sich aus § 137a Abs. 3 Nr. 1 SGB V. Demnach soll das IQTIG unter anderem damit beauftragt werden, „für die Messung und Darstellung der Versorgungsqualität möglichst sektorenübergreifend abgestimmte risikoadjustierte Indikatoren und Instrumente einschließlich Module für Patientenbefragungen auch in digitaler Form zu entwickeln.“

Durch das GVWG wurde die Möglichkeit der Patientenbefragung auch in digitaler Form klargestellt. Die Gesetzesbegründung führt hierzu aus:

„Durch die Ergänzung wird klargestellt, dass das IQTIG bei der Entwicklung von Patientenbefragungen im Rahmen der datengestützten Qualitätssicherung auch digitale Wege zur Erhebung der Daten nutzen soll. Ob eine Patientenbefragung mittels digitaler Medien zu entwickeln ist, dürfte bereits bei der Beauftragung des G-BA nach § 137b Absatz 1 Satz 1 in Abstimmung mit dem IQTIG zu klären sein. Die vorliegende Klarstellung soll die Nutzung digitaler Erhebungsmöglichkeiten unterstützen, zumal das IQTIG und der G-BA bisher die traditionelle Befragung mittels Papierfragebögen verwenden. Welches Medium besser geeignet ist, ist im Einzelfall insbesondere unter Berücksichtigung des konkret zu befragenden Patientenkollektivs zu entscheiden. Für digitale Patientenbefragungen gelten die Vorgaben des § 299 Absatz 4 entsprechend (s. § 299 Absatz 4 Satz 11 neu).“ (BT-Drs 19/26822 S. 95)

Auch der neu eingefügte § 299 Abs. 4 Satz 11 SGB V stellt die Möglichkeit der Onlinebefragung klar. Alle bisherigen Angaben und Regelungen zu Patientenbefragungen gelten uneingeschränkt auch für die digitale Form. Darüber hinaus werden die Voraussetzungen für die Versendestelle gelockert. Sie kann auch andere Aufgaben wahrnehmen.

Da § 299 Abs. 4 SGB V ermöglicht, dass der G-BA „auch die Übermittlung nicht pseudonymisierter personenbezogener Daten der Versicherten und nicht pseudonymisierter personen- oder einrichtungsbezogener Daten der Leistungserbringer vorsehen [kann], soweit dies für die Auswahl der Versicherten oder die Versendung der Fragebögen erforderlich ist“, wäre auch die Übermittlung von Mobilfunk-Rufnummern oder Email-Adressen von Versicherten möglich, um Onlinefragebögen digital zu übermitteln. Da solche personenbezogenen Daten jedoch keine Versichertenstammdaten sind, ist anzunehmen, dass solche Angaben in nur sehr geringem Umfang bei den Leistungserbringern vorliegen. Daher müssten sie gesondert erhoben und ggf. hierzu eine Einwilligung eingeholt werden.

Onlinebefragungen unterliegen somit den gleichen rechtlichen und datenschutzrechtlichen Rahmenbedingungen wie papierbasierte Befragungen. Bei beiden Formen legt der G-BA nach § 299 Abs. 4 SGB V in Richtlinien und Beschlüssen die Einzelheiten des Verfahrens fest.

8 Ergebnisse der Übersicht von systematischen Reviews

Die systematische Recherche in den bibliographischen Datenbanken MEDLINE, PsycInfo, Embase, CINAHL und Cochrane Database of Systematic Reviews (Suchdatum: 1. Februar 2021) sowie den weiteren Informationsquellen (siehe Abschnitt 4.4.2 und 4.4.3) lieferte 14.864 Treffer. Nach Entfernung der Dubletten wurden 9.623 Treffer im Titel- und Abstract gesichtet, wovon 58 ins Volltext-Screening eingeschlossen wurden. Schlussendlich erfüllten 16 Artikel die vorab definierten Einschlusskriterien (siehe Tabelle 2).

Eine Gesamtübersicht über die Recherche nach systematischen Reviews in bibliographischen Datenbanken und weiterer Informationsquellen bietet das nachfolgende PRISMA-Flussdiagramm (Page et al. (2021), Abbildung 15).

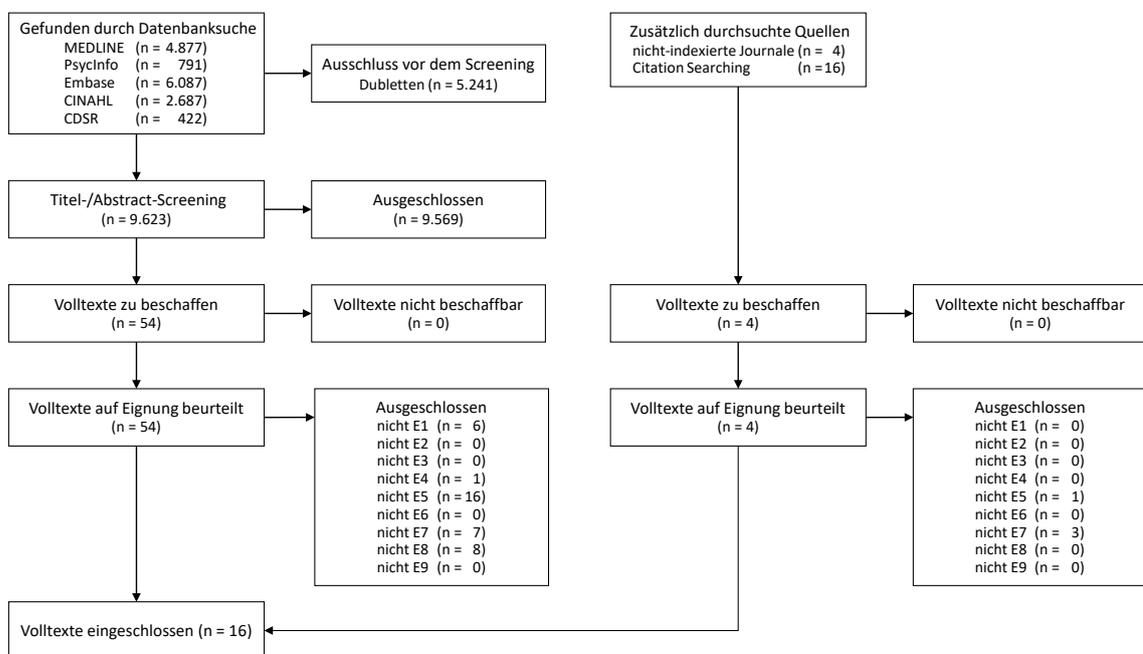


Abbildung 15: Flussdiagramm der Recherche nach systematischen Reviews

8.1 Eingeschlossene Publikationen

Nach dem Volltext-Screening wurden 16 Artikel eingeschlossen (siehe Tabelle 8). Die im Volltext-Screening ausgeschlossenen Artikel finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang 3 (für die Ausschlussgründe vgl. Tabelle 2 in Abschnitt 4.3.1).

Tabelle 8: Eingeschlossene Artikel

	Titel	Referenz
1	Interformat Reliability of Digital Psychiatric Self-Report Questionnaires: A Systematic Review	Alfonsson et al. (2014)
2	Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature	Blumenberg und Barros (2018)
3	Equivalence of electronic and paper-based patient-reported outcome measures	Campbell et al. (2015)
4	Social desirability is the same in offline, online, and paper surveys: A meta-analysis	Dodou und de Winter (2014)
5	Non-randomised evaluations of strategies to increase participant retention in randomised controlled trials: a systematic review	Elfeky et al. (2020)
6	An exploration of the data collection methods utilised with children, teenagers and young people (CTYPs)	Flanagan et al. (2015)
7	Socially Desirable Responding in Web-Based Questionnaires: A Meta-Analytic Review of the Candor Hypothesis	Gnams und Kaspar (2017)
8	Building the Evidence Base for Remote Data Collection in Low- and Middle-Income Countries: Comparing Reliability and Accuracy Across Survey Modalities	Greenleaf et al. (2017)
9	A Comparison of the Types of Screening Tool Administration Methods Used for the Detection of Intimate Partner Violence: A Systematic Review and Meta-Analysis	Hussain et al. (2015)
10	Electronic Data Capture Versus Conventional Data Collection Methods in Clinical Pain Studies: Systematic Review and Meta-Analysis	Jibb et al. (2020)
11	Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods (Review)	Marcano Belisario et al. (2015)
12	Benefits and Disadvantages of Electronic Patient-reported Outcome Measures: Systematic Review	Meirte et al. (2020)
13	Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013	Muehlhausen et al. (2015)
14	Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis	Rutherford et al. (2016)
15	Mean scores for self-report surveys completed using paper-and-pencil and computers: A meta-analytic test of equivalence	Weigold et al. (2018a)
16	A meta-analytic review of measurement equivalence study findings of the SF-36® and SF-12® Health Surveys across electronic modes compared to paper administration	White et al. (2018)

Die 16 eingeschlossenen Arbeiten verteilten sich wie folgt auf die berücksichtigten Publikationsjahre:

- 2014: 2 (Alfonsson et al. 2014, Dodou und de Winter 2014)
- 2015: 5 (Campbell et al. 2015, Flanagan et al. 2015, Hussain et al. 2015, Marcano Belisario et al. 2015, Muehlhausen et al. 2015)
- 2016: 1 (Rutherford et al. 2016)
- 2017: 2 (Gnambs und Kaspar 2017, Greenleaf et al. 2017)
- 2018: 3 (Blumenberg und Barros 2018, Weigold et al. 2018a, White et al. 2018)
- 2020: 3 (Elfeky et al. 2020, Jibb et al. 2020, Meirte et al. 2020)

Je drei systematische Reviews wurden im *Journal of Medical Internet Research* (Alfonsson et al. 2014, Greenleaf et al. 2017, Jibb et al. 2020) bzw. im *Journal Quality of Life Research* publiziert (Campbell et al. 2015, Rutherford et al. 2016, White et al. 2018). Zwei Artikel wurden im *Journal Computers in Human Behavior* veröffentlicht (Dodou und de Winter 2014, Weigold et al. 2018a). Die acht weiteren systematischen Reviews wurden in den Journalen *International Journal of Public Health* (Blumenberg und Barros 2018), *Systematic Reviews* (Elfeky et al. 2020), *BMC: Research Notes* (Flanagan et al. 2015), *Assessment* (Gnambs und Kaspar 2017), *Trauma, Violence & Abuse* (Hussain et al. 2015), *Cochrane Database of Systematic Reviews* (Marcano Belisario et al. 2015), *JMIR Perioperative Medicine* (Meirte et al. 2020) und *Health and Quality of Life Outcomes* (Muehlhausen et al. 2015) publiziert.

Weitere detaillierten Charakteristika der eingeschlossenen systematischen Reviews finden sich in Tabelle A 8 im Anhang.

Systematische Suche

In den systematischen Reviews wurden im Median 4 bibliographische Datenbanken systematisch durchsucht (Range: 1 bis 13). Dabei wurden in 15 von den 16 eingeschlossenen systematischen Reviews 30 verschiedene bibliographische Datenbanken als Informationsquellen benannt. Am häufigsten wurde in MEDLINE (n = 11), Embase (n = 9) und PsycInfo (n = 8) recherchiert. Als weitere Quellen dienten hauptsächlich Google Scholar sowie die Referenzlisten der eingeschlossenen Artikel.

Im überwiegenden Anteil der eingeschlossenen systematischen Reviews (n = 9) wurde ohne zeitliche Einschränkung des Publikationsdatums gesucht (Alfonsson et al. 2014, Blumenberg und Barros 2018, Dodou und de Winter 2014, Flanagan et al. 2015, Hussain et al. 2015, Jibb et al. 2020, Meirte et al. 2020, Weigold et al. 2018a, White et al. 2018). In fünf systematischen Reviews wurde Literatur für einen begrenzten Zeitraum gesucht (Campbell et al. 2015, Elfeky et al. 2020, Marcano Belisario et al. 2015, Muehlhausen et al. 2015, Rutherford et al. 2016). Im aktuellsten systematischen Review von Jibb et al. (2020) wurde bis November 2019 gesucht. Für zwei Arbeiten lagen keine konkreten Informationen hierzu vor (Gnambs und Kaspar 2017, Greenleaf et al. 2017).

Population

Die Mehrzahl der eingeschlossenen Reviews ($n = 11$) betrachtete eine gemischte Population aus gesunden Teilnehmerinnen und Teilnehmern sowie Teilnehmerinnen und Teilnehmern mit gesundheitlichen Einschränkungen. In einer Übersichtsarbeit wurden ausschließlich Studien eingeschlossen, die gesunde Personen befragten (Blumenberg und Barros 2018). In je zwei Arbeiten wurden nur Patientinnen und Patienten befragt (Elfeky et al. 2020, Jibb et al. 2020) bzw. gab es keine Angaben zur untersuchten Population (Gnambs und Kaspar 2017, Greenleaf et al. 2017). Diese Analyse bezog sich teilweise auf die in den Arbeiten genannten Einschlusskriterien sowie den aufgeführten Charakteristika der eingeschlossenen und für die relevanten Endpunkte ausgewerteten Studien.

In neun Übersichtsarbeiten fanden sich in den Einschlusskriterien bzw. den Beschreibungen der Studienpopulation Angaben zum Alter der Teilnehmerinnen und Teilnehmer (Alfonsson et al. 2014, Flanagan et al. 2015, Gnambs und Kaspar 2017, Hussain et al. 2015, Jibb et al. 2020, Marcano Belisario et al. 2015, Meirte et al. 2020, Muehlhausen et al. 2015, White et al. 2018). Davon fanden sich in zwei Publikationen konkrete Angaben über die Auswertung der Daten von Kindern (Marcano Belisario et al. (2015): 4 bis 11 Jahre; Flanagan et al. (2015): ab 8 Jahre).

Fragebögen

Die Autorinnen und Autoren der eingeschlossenen systematischen Reviews definierten selten konkrete Einschränkungen bzgl. der verwendeten Fragebögen. Im Allgemeinen wurden Studien, die validierte und/oder nicht validierte gesundheitsbezogene Fragebögen nutzten, berücksichtigt (Blumenberg und Barros 2018, Greenleaf et al. 2017, Marcano Belisario et al. 2015, Weigold et al. 2018a). Diese waren themen- bzw. krankheitsspezifisch oder auch generischer Art. Lediglich ein systematischer Review fokussierte ausschließlich auf den SF-12 bzw. SF-36 Fragebogen (White et al. 2018). Andere Arbeiten schlossen Studien mit verschiedenen Fragebögen zu einem festgelegten (medizinischen) Thema ein (mentale Gesundheit: Alfonsson et al. 2014, Schmerz: Jibb et al. 2020, häusliche Gewalt: Hussain et al. 2015, soziale Erwünschtheit: Dodou und de Winter 2014, Gnambs und Kaspar 2017). Vier Publikationen waren themenoffen und schränkten lediglich auf (validierte) PROMs ein (Campbell et al. 2015, Meirte et al. 2020, Muehlhausen et al. 2015, Rutherford et al. 2016); zwei weitere machten keine weiteren Angaben zu den gesundheitsbezogenen Fragebögen (Elfeky et al. 2020, Flanagan et al. 2015).

Anzahl und Design der Studien

Unabhängig von den untersuchten Endpunkten schlossen die systematischen Reviews im Median 35 Publikationen ein (Range: 6 bis 159). Die genauen Studiendesigns der in den systematischen Reviews eingeschlossenen Publikationen wurden nur zum Teil berichtet. Dies beinhaltete sowohl Studien im Paralleldesign als auch (randomisierte) Cross-Over-Studien. In lediglich einer Arbeit gab es keine detaillierteren Informationen zum eingeschlossenen kontrollierten Studiendesign (Flanagan et al. 2015). Weitere Vergleiche wie z. B. Kohorten- oder Querschnittstudien

wurden weniger häufig eingeschlossen. Für einzelne Publikationen blieb das Studiendesign unklar, sodass diese Studien nicht zur Beantwortung der Forschungsfrage herangezogen werden konnte.

Potenzielle Interessenkonflikte und Finanzierung

In nahezu allen eingeschlossenen systematischen Reviews wurden Angaben dahingehend gemacht, ob potenzielle Interessenkonflikte vorlagen oder nicht (n = 14). Von diesen wurde in 11 Studien angegeben, dass keine potenziellen Interessenkonflikte vorlagen. In drei Artikeln fanden sich detailliertere Angaben zu potenziellen beruflichen, materiellen oder finanziellen Verbindungen von einem Autor (Campbell et al. 2015, Jibb et al. 2020) oder allen Autorinnen und Autoren (White et al. 2018). Bei dem Letztgenannten handelte es sich um Mitarbeitende eines Unternehmens, das kommerzielle Interessen hinsichtlich des einzigen in diesem systematischen Review eingeschlossenen Fragebogens (SF-12 bzw. SF-36) verfolgte.

Bezüglich der Finanzierung der Durchführung der systematischen Reviews fanden sich in 9 Publikationen Angaben. In drei dieser Artikel wurde berichtet, dass es keine externe Finanzierung der Arbeit gab (Blumenberg und Barros 2018, Elfeky et al. 2020, Gnams und Kaspar 2017). Bei einer Publikation waren die Autorinnen und Autoren Beschäftigte eines kommerziellen Unternehmens (White et al. 2018, siehe oben). In den anderen Fällen wurden entsprechende unabhängige (staatliche bzw. gemeinnützige) Förderungen genannt.

Überlappung

In den 16 systematischen Reviews mit klassischer Metaanalyse und/oder deskriptiver Beschreibung der Daten zu einem hier ausgewerteten Endpunkt wurden (rein additiv) insgesamt 564 Publikationen eingeschlossen (zwischen 1 und 159 Publikationen pro systematischem Review). Aufgrund der Überlappungen, d. h., dass einzelne Publikationen in mehr als einem systematischen Review eingeschlossen waren, betrug die Anzahl der tatsächlich berücksichtigten Publikationen 371. Der überwiegende Teil davon (n = 255; 69 %) wurde nur einmalig in einem systematischen Review eingeschlossen und ausgewertet.

In sechs systematischen Reviews wurde die Response-Rate betrachtet (siehe Tabelle A 10 im Anhang und Abschnitt 8.3.1). In diesen Reviews wurden insgesamt 25 Publikationen eingeschlossen und ausgewertet. Da alle 25 Publikationen nur einmalig genutzt wurden, gab es keine Überlappung. Der CCA-Wert betrug 0,00 %.

Hinsichtlich der Übereinstimmung des Antwortverhaltens wurden 14 systematische Reviews berücksichtigt. In diesen wurden rein additiv 544 Publikationen eingeschlossen (zwischen 2 und 159 Publikationen pro systematischem Review). Da mehrere Publikationen in mehr als einem systematischen Review ausgewertet wurden, betrug die Anzahl der tatsächlich berücksichtigten Publikationen 354. Der überwiegende Teil davon (n = 238; 67 %) wurde nur einmalig in einem systematischen Review eingeschlossen und ausgewertet. Drei Publikationen wurden in sechs systematischen Reviews analysiert (siehe Tabelle 9).

Tabelle 9: Verteilung der Publikationen zur Übereinstimmung des Antwortverhaltens

Häufigkeit einzelner Publikationen	1	2	3	4	5	6
Anzahl der Publikationen	238	70	28	11	4	3

Der korrigierte Überlappungsgrad (CCA) ergab 4,13 %. Somit lag bei dem hier vorliegenden Overview von systematischen Reviews eine sehr geringe Überlappung der eingeschlossenen Publikationen bezüglich der Übereinstimmung des Antwortverhaltens vor. Die systematischen Reviews hatten dabei sowohl breite als auch sehr eingeschränkte Fragestellungen. Auch die verschiedenen betrachteten Populationen und Arten der Datenerhebung, die in den systematischen Reviews untersucht werden sollten und die dadurch bedingten unterschiedlichen Einschlusskriterien, trugen zu dieser geringen Überlappung der Publikationen in den systematischen Reviews bei.

8.2 Kritische Bewertung

Die AMSTAR 2-Bewertungen der einzelnen eingeschlossenen systematischen Reviews sind der Tabelle A 9 im Anhang zu entnehmen. Lediglich zwei Übersichtsarbeiten wurden den Anforderungen des Items 2 (a priori angefertigtes Protokoll) vollständig gerecht (Elfeky et al. 2020, Marcano Belisario et al. 2015) und zwei Reviews teilweise (Jibb et al. 2020, Meirte et al. 2020). Eine umfangreiche systematische Recherche nach Primärliteratur fand dagegen bei der Mehrheit der eingeschlossenen systematischen Reviews statt (Item 4). Einige Reviews wiesen methodische Schwächen bei den berichteten Selektions- und Extraktionsprozessen auf (Item 5: 10/16 und Item 6: 8/16). Weiterhin berichteten nur 6 der 16 Reviews von einer Bewertung ihrer eingeschlossenen Studien (Item 9). Aufgrund der Heterogenität der eingeschlossenen Studien wurde in sieben Reviews der Versuch unternommen, die Daten, die für den hier berichteten Overview relevant waren, metaanalytisch auszuwerten (Dodou und de Winter 2014, Gnambis und Kaspar 2017, Hussain et al. 2015, Muehlhausen et al. 2015, Rutherford et al. 2016, Weigold et al. 2018a, White et al. 2018), wobei in zwei Reviews (Dodou und de Winter 2014, Hussain et al. 2015) die Auswertung als statistisch angemessen eingeschätzt wurde (Item 11).

Für eine Kategorisierung der methodischen Qualität dieser Übersichtsarbeiten wurden sieben kritische Domänen herangezogen. Zwei Reviews verfügten über keine methodischen Schwachstellen in diesen kritischen Domänen (Elfeky et al. 2020, Marcano Belisario et al. 2015). Da bei der Publikation von Marcano Belisario et al. (2015) lediglich eine unkritische Schwachstelle aufgedeckt werden konnte, wurde der Grad des Vertrauens in diese Ergebnisse als *high* eingestuft. Bei Elfeky et al. (2020) gab es dagegen zwei unkritische Schwächen, somit wurde die methodische Qualität dieses systematischen Reviews als *moderate* bewertet.

Weiterhin wurde bei einem systematischen Review die methodische Qualität als *low* eingestuft, da dieser eine methodische Schwachstelle in einer kritischen Domäne aufwies (Meirte et al. 2020). Die übrigen 13 Übersichtsarbeiten besaßen mehr als eine kritische Schwachstelle, weshalb das Vertrauen in die Ergebnisse dieser Reviews mit Hinblick auf die gefundenen Effektgrößen der untersuchten Erhebungsmodi auf Basis des AMSTAR-2-Instruments als *critically low* gewertet wurde (Alfonsson et al. 2014, Blumenberg und Barros 2018, Campbell et al. 2015, Dodou

und de Winter 2014, Flanagan et al. 2015, Gnambs und Kaspar 2017, Greenleaf et al. 2017, Hussain et al. 2015, Jibb et al. 2020, Muehlhausen et al. 2015, Rutherford et al. 2016, Weigold et al. 2018a, White et al. 2018).

8.3 Effekte des Erhebungsmodus

Die Ergebnisse mit Hinblick auf unterschiedliche Response-Raten je nach Erhebungsmodus finden sich in Tabelle A 10 und mit Hinblick auf Unterschiede im Antwortverhalten in Tabelle A 11 und Tabelle A 12 im Anhang.

Die eingeschlossenen systematischen Reviews untersuchten Vergleiche von Paper-Pencil-Befragungen mit einer elektronischen Datenerhebung und verschiedenen elektronischen Datenerhebungen untereinander. Bei der Kategorisierung wurden dabei Synonyme für Paper-Pencil-Befragungen zusammengefasst (*paper and pencil, paper-based tool, postal, etc.*).

Folgende Modi wurden mit Papierfragebögen verglichen:

- Elektronisches Endgerät (allgemein)
- Web/Computer
- Smartphone App
- Tablet
- Palm/ Personal Digital Assistant (PDA)
- Interactive Voice Response System (IVRS)

Des Weiteren fanden sich in den eingeschlossenen systematischen Reviews Daten zu den folgenden elektronischen Erhebungsmodi im Vergleich:

- Smartphone versus Web/Computer
- Tablet versus Web/Computer
- IVRS versus Web/Computer
- Smartphone versus PDA
- Smartphone versus SMS
- PDA versus sonstiger elektronischer Erfassung

Nachfolgend wurden Vergleiche, die Ergebnisse zu den Endpunkten Response-Rate und/oder Antwortverhalten lieferten, im Detail aufgeführt. Bei fehlenden Abschnitten zu einem Vergleich bzw. einem Endpunkt, wurden diese in keinem der eingeschlossenen systematischen Reviews berichtet.

8.3.1 Response-Rate

Fünf identifizierte systematische Reviews lieferten Ergebnisse für den Endpunkt Response-Rate (Blumenberg und Barros 2018, Elfeky et al. 2020, Flanagan et al. 2015, Jibb et al. 2020, Rutherford et al. 2016).

Web/Computer versus Papier

Fünf systematische Reviews untersuchten die Response-Rate einer selbstadministrierten Paper-Pencil-Befragung im Vergleich zur selbstadministrierten Befragung am Computer (Blumenberg und Barros 2018, Elfeky et al. 2020, Flanagan et al. 2015, Jibb et al. 2020, Rutherford et al. 2016).

Ausgehend von der methodischen Qualität berichtete der am besten bewertete systematische Review höhere Response-Raten für eine Befragung mittels Computer im Vergleich zu einer Paper-Pencil-Befragung (Elfeky et al. 2020, moderates Vertrauen in die Ergebnisse). Allerdings beruht dieses Ergebnis auf einer einzigen Primärstudie, in der sich weder der Befragungskontext (12-Monats-Follow-up nach gesundheitsförderlicher Intervention) noch die Zielpopulation (Studentinnen) auf den Kontext der gesetzlichen QS übertragen lassen. Dagegen berichteten Blumenberg und Barros (2018) auf Basis von neun eingeschlossenen Studien und mit der insgesamt größten Anzahl an analysierten Fällen von durchschnittlich höheren Response-Raten bei der Papierversion. Die Autoren weisen allerdings darauf hin, dass aufgrund der erheblichen Heterogenität der Ergebnisse der eingeschlossenen Studien keine validen Schlüsse hinsichtlich der tatsächlichen Größe eines Mode-Effektes im Sinne einer metaanalytischen Evidenz gezogen werden können. Die berechneten Response-Raten der weiteren eingeschlossenen Reviews untermauerten das heterogene Bild (Flanagan et al. 2015, Jibb et al. 2020, Rutherford et al. 2016). So errechneten beispielsweise Flanagan et al. (2015) bei zwei eingeschlossenen Studien mit sehr jungen Probandinnen und Probanden höhere Response-Raten für die Datenerhebung am Computer. Jibb et al. (2020) fanden hingegen bei klinischen Populationen bei zwei eingeschlossenen randomisierten kontrollierten Studien divergierende Ergebnisse: eine Studie fand keinen signifikanten Unterschied, während die zweite Studie eine beinahe doppelt so hohe Response-Rate bei der Verwendung eines Papierfragebogens im Vergleich zum Online-Format feststellte (vgl. 1 in Tabelle A 10 im Anhang).

Smartphone App versus Papierfragebögen

Ein systematischer Review untersuchte die Response-Rate einer selbstadministrierten Paper-Pencil-Befragung im Vergleich zur selbstadministrierten Befragung mit dem Smartphone (Jibb et al. 2020). Die Ergebnisse ließen hier nicht auf unterschiedliche Response-Raten je nach Erhebungsmodus schließen. Stark einschränkend anzumerken ist allerdings, dass die entsprechenden Ergebnisse des Reviews auf einer einzigen Studie basierten, in der sich der Befragungskontext (Erfassung der postoperativen Schmerzstärke auf einer numerischen Bewertungsskala) nicht auf den Kontext der Patientenbefragung in der gesetzlichen QS übertragen lässt.

8.3.2 Antwortverhalten

Vierzehn identifizierte systematische Reviews lieferten Ergebnisse hinsichtlich der Übereinstimmung des Antwortverhaltens zwischen unterschiedlichen selbstadministrierten Befragungsmodi (Alfonsson et al. 2014, Campbell et al. 2015, Dodou und de Winter 2014, Flanagan et al. 2015, Gnams und Kaspar 2017, Greenleaf et al. 2017, Hussain et al. 2015, Jibb et al. 2020, Marcano Belisario et al. 2015, Meirte et al. 2020, Muehlhausen et al. 2015, Rutherford et al. 2016, Weigold et al. 2018a, White et al. 2018).

8.3.2.1 Vergleiche mit Paper-Pencil-Befragung

Alle 14 eingeschlossenen systematischen Reviews, die Informationen für die Übereinstimmung des Antwortverhaltens lieferten, untersuchten auch mindestens eine Art der Paper-Pencil-Befragung im Vergleich (Alfonsson et al. 2014, Campbell et al. 2015, Dodou und de Winter 2014, Flanagan et al. 2015, Gnams und Kaspar 2017, Greenleaf et al. 2017, Hussain et al. 2015, Jibb et al. 2020, Marcano Belisario et al. 2015, Meirte et al. 2020, Muehlhausen et al. 2015, Rutherford et al. 2016, Weigold et al. 2018a, White et al. 2018).

Elektronisches Endgerät versus Papier

Vier systematische Reviews untersuchten die Übereinstimmung des Antwortverhaltens bei selbstadministrierten Paper-Pencil-Befragungen mit dem Antwortverhalten in allgemeinen elektronischen Befragungen (Campbell et al. 2015, Rutherford et al. 2016, Weigold et al. 2018a, White et al. 2018). Die elektronischen Geräte umfassten dabei, soweit angegeben: Computer, Tablet, PDA und Handheld. Unabhängig von der genauen Definition des Endpunktes (*equivalence* (Campbell et al. 2015, Weigold et al. 2018a), *agreement* (White et al. 2018) oder *difference in mean score* (Rutherford et al. 2016)) und der Bewertung der Ergebnisse kamen alle Übersichten zu dem Schluss, dass es keine Modi-assoziierten Unterschiede im Antwortverhalten gab (vgl. 1 in Tabelle A 11 im Anhang). Anzumerken ist, dass Weigold et al. (2018a) eine hohe Heterogenität hinsichtlich des Moduseffekts in Studien mit Between-Group-Design feststellten. Im Gegensatz dazu war die Heterogenität der Ergebnisse aus den eingeschlossenen Studien mit Within-Group Design gering (Weigold 2018a).

Web/Computer versus Papier

Acht systematische Reviews untersuchten Unterschiede im Antwortverhalten bei selbstadministrierten Paper-Pencil-Befragungen im Vergleich zu selbstadministrierten computer- oder webgestützten Befragungen (Alfonsson et al. 2014, Dodou und de Winter 2014, Flanagan et al. 2015, Gnams und Kaspar 2017, Hussain et al. 2015, Jibb et al. 2020, Meirte et al. 2020, Muehlhausen et al. 2015).

Analysierte Endpunkte waren u. a. *interformat reliability*, *agreement* bzw. *equivalence* (berechnet mittels verschiedener Korrelationsmaße oder standardisierter Mittelwertdifferenzen), *social desirability* (als standardisierte Mittelwertdifferenzen) sowie unterschiedliche Häufigkeiten hinsichtlich gemachter Angaben zu Risikoverhalten und Substanzgebrauch. Einzelne Vergleiche basierten auf einer heterogenen Studienbasis. So zeigten sich beispielsweise bei Muehlhausen et al. (2015) deutliche Unterschiede in den 24 eingeschlossenen Studien hinsichtlich der Stärke des Effekts der untersuchten Erhebungsmodi auf das Antwortverhalten. Des Weiteren erschwerte eine unzureichende Berichterstattung der Review-Autorinnen und -Autoren die Interpretation einzelner Ergebnisse. So konnte bei Gnams und Kaspar (2017) hinsichtlich der drei berechneten Meta-Analysen nicht festgestellt werden, ob jeweils die Mehrheit der Teilnehmerinnen und Teilnehmer die Befragung in Anwesenheit einer Beobachterin bzw. eines Beobachters ausgefüllt hat. Diese Ergebnisse waren daher nur eingeschränkt auf die hier zu beantwortende Forschungs-

frage übertragbar. Unterschiedliche und inkonsistente Endpunkt-Definitionen und Beschreibungen eines Teils der Ergebnisse von Meirte et al. (2020) verhinderten wiederum deren Extraktion und Analyse für diese Recherche.

In der überwiegenden Mehrheit der Studien konnte kein Unterschied im Antwortverhalten bei der Beantwortung einer selbstadministrierten Paper-Pencil-Befragung im Vergleich zu einer selbstadministrierten computer- oder webgestützten Befragung aufgezeigt werden (vgl. 2 in Tabelle A 11 im Anhang). Lediglich in einzelnen Reviews, die auf Befragungen zu potenziell sensiblen Themen fokussierten, gab es teilweise Tendenzen, die auf Moduseffekte hindeuteten (siehe Flanagan et al. (2015), Hussain et al. (2015) und Dodou und de Winter (2014)). Alfonsson et al. (2014) fanden in einzelnen Studien darüber hinaus relativ niedrige Werte hinsichtlich der *interformat reliability*. Jibb et al. (2020) interpretierten die Ergebnisse der eingeschlossenen Studien zu Moduseffekten bei der Messung von selbstberichteten Schmerzen zum Großteil als *equivalent*, die Ergebnisse einzelner Studien aber auch als *non-equivalent* (1 Studie) sowie *discrepant* (2 Studien).

Smartphone versus Papier

Drei systematische Reviews untersuchten die Übereinstimmung des Antwortverhaltens bei selbstadministrierten Paper-Pencil-Befragungen mit selbstadministrierten Befragungen mittels Smartphone (meist als App; Alfonsson et al. 2014, Jibb et al. 2020, Marcano Belisario et al. 2015).

Die systematischen Reviews von Alfonsson et al. (2014) sowie Jibb et al. (2020) zeigten bzgl. *interformat reliability* und *pain score equivalence* keine Unterschiede in der Übereinstimmung des Antwortverhaltens bei der selbstadministrierten Beantwortung des Fragebogens auf Papier bzw. mit dem Smartphone. Die Ergebnisse zur Datenäquivalenz des in diesem Overview eingeschlossenen systematischen Cochrane-Reviews von Marcano Belisario et al. (2015) zeigte ebenfalls keine Unterschiede bei der Beantwortung einer selbstadministrierten Befragung mittels Papier- oder App-Version auf dem Smartphone (vgl. 3 in Tabelle A 11 im Anhang).

Tablet versus Papier

Drei systematische Reviews untersuchten die Übereinstimmung des Antwortverhaltens von selbstadministrierten Paper-Pencil-Befragungen mit dem Antwortverhalten in selbstadministrierten Befragungen mittels Tablet (Jibb et al. 2020, Meirte et al. 2020, Muehlhausen et al. 2015).

Die in den systematischen Reviews untersuchten Vergleiche analysierten die Endpunkte *score equivalence*, *agreement* und *correlation*. Alle Ergebnisse deuten auf eine hohe Übereinstimmung des Antwortverhaltens zwischen einer Paper-Pencil- und Tablet-Befragung hin. Unterschiedliche und inkonsistente Endpunkt-Definitionen und Beschreibungen eines Teils der Ergebnisse von Meirte et al. (2020) bzgl. *data completion*, *missing data* und *completion rate* verhinderten deren Extraktion und Analyse für diesen Overview (vgl. 4 in Tabelle A 11 im Anhang).

Palm/PDA versus Papier

Vier systematische Reviews untersuchten die Übereinstimmung des Antwortverhaltens einer selbstadministrierten Paper-Pencil-Befragung im Vergleich zu einer selbstadministrierten Befragung mittels Palm/Personal Digital Assistant (PDA; Alfonsson et al. 2014, Jibb et al. 2020, Meirte et al. 2020, Muehlhausen et al. 2015).

Die in den systematischen Reviews untersuchten Vergleiche analysierten die Endpunkte *interformat reliability, score equivalence, agreement* und *correlation*. Alle Ergebnisse zeigten keinen Unterschied bzgl. der Übereinstimmung des Antwortverhaltens zwischen Paper-Pencil- und Palm/PDA-Befragung. Meirte et al. (2020) präsentierten einzelne Ergebnisse unklar, sodass eine Interpretation für diesen Overview teilweise nicht möglich war (vgl. 5 in Tabelle A 11 im Anhang).

IVRS versus Papier

Die beiden eingeschlossenen systematischen Reviews von Greenleaf et al. (2017) und Muehlhausen et al. (2015) untersuchten Unterschiede zwischen dem Antwortverhalten in selbstadministrierten Paper-Pencil-Befragungen und dem Antwortverhalten in Befragung mittels *Interactive Voice Response System* (IVRS), d. h. einer computergestützten telefonischen Befragung.

Untersucht wurden die Endpunkte *item non-response rate, correlation* und *rate of socially desirable answers*. Fokus der Untersuchung von Greenleaf et al. (2017) war die Datenerhebung in Ländern mit niedrigem und mittlerem Einkommen. In einer eingeschlossenen randomisierten kontrollierten Studie war die *rate of socially desirable answers* in der Gruppe der postalischen Paper-Pencil-Befragung höher als bei IVRS. Hinsichtlich der *item non-response rate* wurde von einem statistisch signifikanten Unterschied ohne weitere Details berichtet. Muehlhausen et al. (2015) berichtete bei fünf berücksichtigten Studien von einer sehr guten Korrelation zwischen Paper-Pencil- und IVRS-Befragung (vgl. 6 in Tabelle A 11 im Anhang).

8.3.2.2 Vergleiche elektronischer Erhebungsmodi

Vier eingeschlossene systematische Reviews, die Informationen hinsichtlich der Übereinstimmung des Antwortverhaltens lieferten, untersuchten Vergleiche verschiedener elektronischer Möglichkeiten einer selbstadministrierten Befragung (Greenleaf et al. 2017, Jibb et al. 2020, Marcano Belisario et al. 2015, Muehlhausen et al. 2015).

Smartphone App versus Web/Computer

Zwei systematische Reviews untersuchten das Antwortverhalten von selbstadministrierten Befragungen ausgefüllt am Smartphone (web-basiert) oder mittels App (iOS) im Vergleich zur Beantwortung am Computer (web-basiert) oder mittels Laptop (Greenleaf et al. 2017, Marcano Belisario et al. 2015).

Die Ergebnisse der Endpunkte *rate of reporting of sensitive behavior* und *data equivalence* zeigten im Allgemeinen keinen Unterschied bei den beiden Erhebungsmodi in Bezug auf die gemessene Übereinstimmung des Antwortverhaltens (vgl. 1 in Tabelle A 12 im Anhang).

Tablet versus Web/Computer

Ein systematischer Review von Muehlhausen et al. (2015) untersuchte die Übereinstimmung des Antwortverhaltens zwischen selbstadministrierten Befragungen mittels Tablet und dem Antwortverhalten am Computer (web-basiert).

Berechnet wurde die Korrelation, deren Interpretation eine sehr hohe Übereinstimmung beider Erhebungsmodi in Bezug auf die gemessene Übereinstimmung des Antwortverhaltens zeigte (vgl. 2 in Tabelle A 12 im Anhang).

IVRS versus Web/Computer

Ein systematischer Review von Muehlhausen et al. (2015) untersuchte die Übereinstimmung des Antwortverhaltens einer selbstadministrierten Befragung mittels *Interactive Voice Response System* (IVRS), d. h. einer computergestützten telefonischen Befragung im Vergleich zum Computer (web-basiert).

In der eingeschlossenen Studie wurde eine Korrelation berechnet, deren Interpretation eine sehr hohe Übereinstimmung beider Erhebungsmodi in Bezug auf die gemessene Übereinstimmung des Antwortverhaltens zeigte (vgl. 3 in Tabelle A 12 im Anhang).

Smartphone-App versus PDA

Ein systematischer Review von Marcano Belisario et al. (2015) untersuchte die Übereinstimmung des Antwortverhaltens einer selbstadministrierten Befragung mittels Smartphone App (iOS) im Vergleich zu einem Personal Digital Assistant (PDA).

Die Autoren des Cochrane Reviews interpretierten die Ergebnisse für den Vergleich dieser Erhebungsmodi als *equivalent* (vgl. 4 in Tabelle A 12 im Anhang).

Smartphone-App versus SMS

Ein systematischer Review von Marcano Belisario et al. (2015) untersuchte die Übereinstimmung des Antwortverhaltens einer selbstadministrierten Befragung mittels Smartphone-App im Vergleich zu einer Befragung mittels SMS.

In einer Studie mit relativ kleiner Fallzahl konnte bzgl. des Endpunkts *data equivalence* kein statistisch signifikanter Unterschied ermittelt werden. Im Gegensatz dazu wurde für den Endpunkt *data completeness* in derselben Studie ein statistisch signifikanter Unterschied ermittelt. Lediglich an Tag 3 der 6-tägigen Befragung wurde kein statistisch signifikanter Unterschied gemessen (vgl. 5 in Tabelle A 12 im Anhang).

PDA versus sonstige elektronische Erfassung

Ein systematischer Review von Jibb et al. (2020) untersuchte die Übereinstimmung des Antwortverhaltens einer selbstadministrierten Befragung mittels Personal Digital Assistant (PDA) im Vergleich zu einer anderen Art der elektronischen Datenerhebung (*manually manipulated slide device-based tool*) ohne detailliertere Beschreibung. Das Ergebnis der einen eingeschlossenen Studie wurde von den Autoren mit *equivalent* interpretiert (vgl. 6 in Tabelle A 12 im Anhang).

8.4 Zusammenfassung

In der vorliegenden Übersicht von systematischen Reviews konnten 16 systematische Übersichtsarbeiten identifiziert und eingeschlossen werden, die Informationen zur Beantwortung der vorab spezifizierten Fragestellungen lieferten. Die Populationen in den eingeschlossenen Artikeln war heterogen, d. h. neben nicht klinischen Probandinnen und Probanden wurde ein breites Spektrum an Patientinnen und Patienten, z. B. mit einer Krebserkrankung, Schmerzen oder einer psychischen Erkrankung bzw. unterschiedlichen Alters eingeschlossen. Ebenfalls heterogen waren die eingeschlossenen Übersichtsarbeiten hinsichtlich der untersuchten selbstadministrierten gesundheitsbezogenen Fragebögen. Nur in einem systematischen Review lag der Fokus auf einem bestimmten Instrument.

Etwaige potenzielle Interessenkonflikte von Autorinnen und Autoren der eingeschlossenen systematischen Reviews hatten entsprechend der Angaben keinen Einfluss auf deren Untersuchung. Darüber hinaus war für die Interpretation der Ergebnisse des Overviews eine eventuelle Überlappung der in den systematischen Reviews eingeschlossenen Publikationen vernachlässigbar. Der überwiegende Teil der eingeschlossenen systematischen Reviews wies nach Anwendung des modifizierten AMSTAR-2-Instruments eine sehr geringe methodische Qualität auf, was das Vertrauen in die berichteten Ergebnisse insbesondere mit Hinblick auf konkrete und eindeutig interpretierbare Effektstärken einschränkt. Drei Reviews erzielten eine bessere methodische Bewertung und ermöglichten so eine vertrauensvollere Einordnung der Ergebnisse.

Die fünf systematischen Reviews, die hinsichtlich des Endpunkts *Response-Rate* eingeschlossen werden konnten, untersuchten zwei unterschiedliche Vergleiche mit Papierfragebögen: 1) Web/Computer und 2) Smartphone-App. Die Ergebnisse waren heterogen, wobei die Richtung des Einflusses des Erhebungsmodus neben anderen Faktoren (z. B. Kontext der Befragung) mit der Zielpopulation assoziiert zu sein schien. So deuten vereinzelte Arbeiten darauf hin, dass insbesondere sehr junge, nicht klinische Bevölkerungsgruppen besser mit online- bzw. computerbasierten Befragungen zu erreichen sind, als mit Papierfragebögen. Demgegenüber steht das Ergebnis eines breit angelegten Reviews zu Befragungen im Public-Health-Bereich, dass onlinebasierte Befragungen generell mit niedrigeren Rücklaufquoten assoziiert sind. Weitere Studien fanden wiederum divergierende bzw. keine Unterschiede der Response-Raten zwischen diesen Erhebungsmodi, wobei letzteres auch auf den in einer einzigen Studie untersuchten Vergleich zwischen Smartphone-App und Papierfragebögen zutraf.

Für den Endpunkt *Antwortverhalten* konnten 14 systematische Reviews eingeschlossen werden. Dabei gab es sieben Vergleiche mit papierbasierten Befragungen: 1. elektronische Endgeräte allgemein, 2. Web/Computer, 3. Smartphone, 4. Tablet, 5. Palm/PDA, 6. IVRS und 7. der Ort der Paper-Pencil-Befragung (Klinik oder zu Hause). Außerdem wurden in vier systematischen Reviews sechs verschiedene elektronische Modi untereinander verglichen: 1. Smartphone-App versus Web/Computer, 2. Tablet versus Web/Computer, 3. IVRS versus Web/Computer, 4. Smartphone-App versus PDA, 5. Smartphone App versus SMS, 6. PDA versus sonstige elektronische Erfassung. Insgesamt zeigten sich keine Unterschiede im Antwortverhalten zwischen verschiedenen elektronischen Endgeräten im Vergleich mit selbstadministrierten Paper-Pencil-

Befragungen. Auch die Ergebnisse der Vergleiche einzelner elektronischer Erhebungsmodi deuteten auf eine Modus-übergreifende Übereinstimmung des Antwortverhaltens hin.

9 Ergebnisse der empirischen Untersuchung

9.1 Bruttostichprobenbeschreibung

Die Befragung fand vom 07. Mai 2021 bis zum 25. Juni 2021 statt. Für die Rekrutierung der benötigten Patientinnen und Patienten zur empirischen Testung von Onlinebefragungen wurden 15 Dialyseeinrichtungen, die sich vorab zur Unterstützung bereit erklärten, anhand bestimmter Kriterien ausgewählt (siehe Abschnitt 5.6).

Insgesamt konnten n = 435 Patientinnen und Patienten rekrutiert werden, von denen 15 die Einverständniserklärung nicht vollständig ausgefüllt hatten, womit sich eine Bruttostichprobe von n = 420 ergab. Nach Bereinigung dieser Bruttostichprobe (siehe Tabelle 10), d. h. nach Abzug der Patientinnen und Patienten, die ihre Einwilligung widerriefen (n = 4), verstorben waren (n = 1) oder deren Anschreiben nicht zustellbar war (n = 10), belief sich die bereinigte Bruttostichprobe auf n = 405 Patientinnen und Patienten. Davon befanden sich schlussendlich n = 101 im Kontrollarm „Single Mode Paper“, n = 102 im Arm „Single Mode Web“, n = 102 im Arm „Mixed-Mode“ und n = 100 im Arm „Push-to-Web“.

Tabelle 10: Überblick über die Bruttostichprobe

Stichprobe	Gesamt	Single Mode Paper	Single Mode Web	Mixed-Mode	Push-to-Web
unbereinigte Bruttostichprobe	435				
unvollständige Einverständniserklärungen	15				
versendet	420	105	105	105	105
unzustellbar	10	3	3	3	1
Teilnahme widerrufen	4	0	0	0	4
verstorben	1	1	0	0	0
bereinigte Bruttostichprobe	405	101	102	102	100

Die Patientinnen und Patienten wurden, wie in Abschnitt 5.4 beschrieben, vor der Feldphase proportional nach den Altersquartilen geschichtet und per Zufallsauswahl den vier experimentellen Armen zugeordnet. Ein Überblick über die Charakteristika der Patientinnen und Patienten je experimentellem Arm, basierend auf den Informationen der Leistungserbringerdaten (Geschlecht, Alter sowie Modalität und Organisationsform der Dialyse), bietet Tabelle 11. Hier ist erkennbar, dass durch die proportionale Schichtung nach Alter der Mittelwert und Median in allen experimentellen Armen nahezu identisch ist. Die Patientinnen und Patienten waren im

Durchschnitt ca. 66 Jahre alt.²⁵ Auch innerhalb der anderen Variablen sind die Verteilungen sehr ähnlich. Lediglich bezüglich des Geschlechts weist der Arm „Single Mode Paper“ laut Chi²-Test signifikant (jeweils $p \leq 0,05$) weniger Frauen auf, als die anderen Arme, was aufgrund der zum Zeitpunkt der Einwilligung zur Studienteilnahme nicht bekannten Zuteilung zu einer bestimmten Studienbedingung als zufälliger Effekt gewertet wurde. Bezüglich der Dialysemodalität wurde über alle Arme verteilt überwiegend Hämodialyse durchgeführt, weshalb fast ausschließlich Daten aus Fragebogen zur Dialyse in einer Einrichtung (Fragebogen A-e) zur Auswertung herangezogen werden konnten.

Tabelle 11: Zusammensetzung der bereinigten Bruttostichprobe

Charakteristika	Single Mode Paper	Single Mode Web	Mixed-Mode	Push-to-Web
Geschlecht				
Weiblich; n (%)	29 (28,7)	46 (45,1)	46 (45,1)	45 (45)
Männlich; n (%)	72 (71,3)	56 (54,9)	56 (54,9)	55 (55)
Alter in Jahren				
Med. (10. und 90. Perzentil)	67 (45; 82)	65 (51; 83,9)	66 (48; 83)	68 (44; 84)
MW (SD)	65,5 (14,1)	66,3 (12,9)	65,3 (14,2)	65,8 (15,4)
Frühes Erwachsenenalter (19–34); n (%)	2 (2)	1 (1)	3 (2,9)	4 (4)
Mittleres Erwachsenenalter (35–64); n (%)	44 (43,6)	49 (48)	44 (43,1)	39 (39)
Höheres Erwachsenenalter (65–79); n (%)	38 (37,6)	30 (29,4)	38 (37,3)	36 (36)
Hohes Alter (80 und älter); n (%)	17 (16,8)	22 (21,6)	17 (16,7)	21 (21)
Derzeitige Dialysemodalität				
HD; n (%)	99 (98)	101 (99)	95 (93,1)	97 (97)
PD; n (%)	2 (2)	1 (1)	7 (6,9)	3 (3)
Fragebogen				

²⁵ In einer linearen Regression mit Alter als abhängige und Einladungsform bzw. experimentellen Arm als unabhängiger Variable (mit „Single Mode Paper“ als Referenzkategorie) wurde zudem getestet, ob die Einladungsform einen signifikanten Einfluss auf das Alter hat. Es konnte mittels F-Test ($p = 0,372$) des Modells gezeigt werden, dass die Einladungsform keinen signifikanten Effekt auf das Alter hatte.

Charakteristika	Single Mode Paper	Single Mode Web	Mixed-Mode	Push-to-Web
Dialyse in einer Einrichtung; n (%)	100 (99)	101 (99)	98 (96,1)	97 (97)
Heimdialyse; n (%)	1 (1)	1 (1)	4 (3,9)	3 (3)
Gesamt; n (%)	101 (100)	102 (100)	102 (100)	100 (100)

Abkürzungen: Med. = Median; MW = Mittelwert; SD = Standardabweichung; n = Anzahl; HD = Hämodialyse; PD = Peritonealdialyse

9.2 Analysen zum Fragebogenrücklauf und zur Nonresponse

9.2.1 Rücklaufquote

Von den n = 420 Patientinnen und Patienten der unbereinigten Bruttostichprobe schickten insgesamt n = 314 einen ausgefüllten Fragebogen zurück.²⁶ Damit ergibt sich eine Gesamtrücklaufquote nach AAPOR RR6 (AAPOR 2016) von 75 %.²⁷ Die Gesamtrücklaufquote auf Basis der bereinigten Bruttostichprobe ergab 77 %. Anzumerken ist, dass einige Leistungserbringer bei der Rekrutierung berichteten, dass es aufgrund der gegebenen Zielsetzung der Studie teilweise schwierig gewesen sei, Patientinnen und Patienten ohne Internetanschluss zur Teilnahme zu motivieren. Es scheint daher wahrscheinlich, dass sich in der Stichprobe überproportional Internet-affine Patientinnen und Patienten befanden.

Die Rücklaufquoten in den einzelnen Studienarmen auf Basis der bereinigten Bruttostichprobe unterschieden sich deutlich, mit 91 % (n = 92) bei „Single Mode Paper“, 60 % (n = 61) bei „Single Mode Web“, 84 % (n = 86) bei „Mixed-Mode“ und 73 % (n = 73) bei „Push-to-Web“. Einen Überblick über die Rücklaufquoten, aufgeteilt nach den experimentellen Armen bzw. Einladungsformen, bietet Tabelle 12. Zusätzlich werden hier die p-Werte der Chi²-Tests berichtet sowie das Effektstärkemaß Cramers V, mit welchem getestet wurde, ob und wie stark sich die Rücklaufquoten der einzelnen experimentellen Arme vom Kontrollarm „Single Mode Paper“ unterscheiden.

Der Arm „Single Mode Web“ hat mit einer Rücklaufquote von 60 % den niedrigsten Rücklauf. Zudem ist die Response-Rate in diesem Arm signifikant niedriger als im Kontrollarm „Single Mode Paper“, was durch einen mittelstarken Effekt bei einem Cramers V von 0,36 verdeutlicht wird. Auch „Push-to-Web“ weist eine signifikant niedrigere Rücklaufquote von 73 % auf, hat jedoch mit einem Cramers V von 0,24 einen schwachen Effekt. Der experimentelle Arm „Mixed-Mode“ hat zwar mit 84 % eine niedrigere Rücklaufquote im Vergleich zu „Single Mode Paper“, allerdings ist der Unterschied nicht signifikant. Folglich kann davon ausgegangen werden, dass

²⁶ Mit „zurückgeschickten Fragebögen“ sind auch jene Fragebögen gemeint, die online ausgefüllt wurden.

²⁷ Die „Standard Definitions“ der American Association for Public Opinion Research (AAPOR) geben 6 Definitionen für das Berichten von Response-Raten vor. Die Nr. 6 der Response-Rate (RR6) ist laut AAPOR definiert als: $RR6 = (I+P) / (I+P) + (R+NC+O)$, wobei I als vollständige Interviews, P als partielle Interviews, R als Verweigerungen, NC als nicht Erreichte und O als andere Gründe definiert sind.

sich „Mixed-Mode“ und „Single Mode Paper“ in ihren Rücklaufquoten wahrscheinlich nicht unterscheiden.

Tabelle 12: Rücklaufquote nach Einladungsform

Einladungsform und Erhebungsmodus	Rücklaufquote	p-Wert	Cramers V
Single Mode Paper	91 % (n = 92)		
Single Mode Web	60 % (n = 61)	p < 0,001	0,36
Mixed-Mode	84 % (n = 86)	p = 0,141	0,10
Push-to-Web	73 % (n = 73)	p < 0,001	0,24

Ferner wurde im Rahmen einer Unit-Nonresponse-Analyse untersucht, ob sich die Befragten, die einen Fragebogen zurückgeschickt haben, in Bezug auf ihr Alter und ihr Geschlecht von den Befragten unterscheiden, die keinen Fragebogen zurückgeschickt haben. Die Analysen, dargestellt in Abbildung A 1 im Anhang, lassen nicht den Schluss zu, dass bezüglich der Variablen Alter und Geschlecht eine systematische Verzerrung durch Unit-Nonresponse vorliegt.

Rücklauf in zeitlichem Bezug zu den Erinnerungsschreiben

Jeweils im Abstand von zwei Wochen wurde nach dem Erstversand der Fragebögen bzw. der Einladung zur Teilnahme an der Onlinebefragung ein Erinnerungsschreiben verschickt. Tabelle 13 zeigt den Anteil der Fragebögen in Abhängigkeit der Einladungsform bzw. des experimentellen Arms, die vor und nach den jeweiligen Erinnerungsschreiben eingegangen waren. Die Erinnerungsschreiben wurden immer an einem Freitag versendet. Für den Versand der Erinnerungsschreiben wurden zwei Werktage angenommen, und da an einem Sonntag keine Post ausgeliefert wird, wurde für die Darstellung des Rücklaufs in Bezug auf die Erinnerungsschreiben statt 14 Tagen zeitlicher Differenz eine Differenz von 17 Tagen berücksichtigt.

Tabelle 13: Rücklauf in zeitlichem Bezug zu den Erinnerungsschreiben

Eingang des Fragebogens...	... vor 1. Erinnerung	... zwischen 1. und 2. Erinnerung	... nach 2. Erinnerung
Single Mode Paper; n (%)	54 (58,7)	28 (30,4)	10 (10,9)
Single Mode Web; n (%)	51 (83,6)	7 (11,5)	3 (4,9)
Mixed-Mode; n (%)	60 (69,8)	20 (23,3)	6 (7,0)
Push-to-web; n (%)	34 (46,6)	11 (15,1)	28 (38,4)

Es zeigt sich, dass der Großteil der Fragebögen vor dem ersten Erinnerungsschreiben zurückgeschickt wurde. Lediglich im experimentellen Arm „Push-to-Web“ wurde etwas weniger als die Hälfte der Fragebögen vor dem ersten Erinnerungsschreiben zurückgesendet. Allerdings ist in

dieser Gruppe zu erkennen, dass ca. 38 % (n = 28) der Respondentinnen und Respondenten erst nach dem zweiten Erinnerungsschreiben geantwortet haben. Von diesen Respondentinnen und Respondenten haben alle den Fragebogen in Papierform ausgefüllt, was deutlich macht, dass das Angebot des Moduswechsels von der Online- in die Papierform deutlichen zusätzlichen Rücklauf generiert hat. Am geringsten ist der Effekt der Erinnerungsschreiben in der Single-Mode-Web-Gruppe ausgeprägt. Hier konnten durch erstes und zweites Erinnerungsschreiben lediglich 10 weitere Teilnehmerinnen und Teilnehmer gewonnen werden.

9.2.2 Beschreibung der Nettostichprobe

Tabelle 14 zeigt die Charakteristika der Patientinnen und Patienten, die einen Fragebogen zurückgeschickt haben bzw. einen Onlinefragebogen ausgefüllt haben. Über alle experimentellen Arme verteilt, lag der Anteil der Männer bei über 50 %. Der Altersdurchschnitt war in allen Gruppen ähnlich und lag bei ca. 65 Jahren. Beim überwiegenden Anteil der Patientinnen und Patienten erfolgte die Dialyse in einer Einrichtung, nur 9 Personen erhielten eine Heimdialyse. Hinsichtlich des Geschlecht ist die Verteilung über die 4 Arme vergleichbar mit der Verteilung in der Bruttostichprobe.

Tabelle 14: Beschreibung der Nettostichprobe

Charakteristika	Single Mode Paper	Single Mode Web	Mixed-Mode	Push-to-Web
Modus				
Online; n (%)	0 (0)	61 (100)	13 (15,1)	45 (61,6)
Papier; n (%)	92 (100)	0 (0)	73 (84,9)	28 (38,4)
Endgerät				
Desktop; n (%)	0 (0)	28 (45,9)	7 (53,8)	21 (46,7)
Mobiltelefon; n (%)	0 (0)	29 (47,5)	5 (38,5)	20 (44,4)
Tablet; n (%)	0 (0)	4 (6,6)	1 (7,7)	4 (8,9)
Geschlecht				
Weiblich; n (%)	27 (29,7)	27 (45,8)	36 (42,9)	31 (43,1)
Männlich; n (%)	64 (70,3)	32 (54,2)	47 (56)	40 (55,6)
Divers; n (%)	0 (0)	0 (0)	1 (1,2)	1 (1,4)
Alter in Jahren				

Charakteristika	Single Mode Paper	Single Mode Web	Mixed-Mode	Push-to-Web
Med. (10. und 90. Perzentil)	67,4 (44,9; 82,2)	63,6 (46,4; 80,5)	68,1 (50,1; 83,3)	65,3 (44,0; 82,3)
MW (SD)	65,2 (14,2)	63,9 (13,3)	67,5 (13,3)	64,1 (15,3)
Frühes Erwachsenenalter (19–34); n (%)	2 (2,2)	1 (1,6)	1 (1,2)	4 (5,6)
Mittleres Erwachsenenalter (35–64); n (%)	40 (43,5)	30 (49,2)	28 (33,7)	30 (41,7)
Höheres Erwachsenenalter (65–79); n (%)	32 (34,8)	20 (32,8)	35 (42,2)	24 (33,3)
Hohes Alter (80 und älter); n (%)	18 (19,6)	10 (16,4)	19 (22,9)	14 (19,4)
Derzeitige Dialysemodalität				
HD; n (%)	82 (96,5)	60 (98,4)	73 (90,1)	66 (95,7)
PD; n (%)	3 (3,5)	1 (1,6)	8 (9,9)	3 (4,3)
Fragebogen				
Dialyse in einer Einrichtung; n (%)	91 (98,9)	60 (98,4)	82 (95,3)	70 (95,9)
Heimdialyse; n (%)	1 (1,1)	1 (1,6)	4 (4,7)	3 (4,1)
Höchster schulischer Bildungsabschluss*				
Noch in schulischer Ausbildung; n (%)	2 (2,4)	1 (1,6)	1 (1,3)	2 (3)
Niedriger schulischer	31 (37,8)	33 (54,1)	38 (50)	29 (43,9)

Charakteristika	Single Mode Paper	Single Mode Web	Mixed-Mode	Push-to-Web
Abschluss; n (%)				
Mittlerer schulischer Abschluss; n (%)	31 (37,8)	14 (23)	23 (28,9)	21 (31,8)
Hoher schu- lischer Ab- schluss; n (%)	16 (19,5)	11 (18)	14 (18,4)	13 (19,7)
Anderer schulischer Abschluss; n (%)	2 (2,4)	2 (3,3)	1 (1,3)	1 (1,5)
Komorbiditäten				
Med. (10. und 90. Perzentil)	3 (1–6)	3 (1–5)	3 (1–6)	3 (1–5)
MW (SD)	3,4 (1,7)	3,2 (1,8)	3,3 (1,8)	2,9 (1,7)
Gesamt; n (%)	92	61	86	73

Abkürzungen: Med. = Median; MW = Mittelwert; SD = Standardabweichung; n = Anzahl; HD = Hämodialyse; PD = Peritonealdialyse

* Schulischer Bildungsabschluss wurde in drei Kategorien aufgeteilt: *niedriger schulischer Abschluss*: von der Schule abgegangen ohne Abschluss / Haupt- oder Volksschulabschluss / Abschluss der Polytechnischen Oberschule der DDR, 8. oder 9. Klasse / noch Schülerin bzw. Schüler einer allgemeinbildenden Vollzeitschule; *mittlerer schulischer Abschluss*: Realschulabschluss / Abschluss der Polytechnischen Oberschule der DDR, 10. Klasse; *hoher schulischer Abschluss*: Fachhochschulreife, Abschluss einer Fachoberschule / allgemeine oder fachgebundene Hochschulreife/Abitur (Gymnasium bzw. EOS, auch EOS mit Lehre).

9.2.3 Moduspräferenzen in den Studienarmen „Mixed-Mode“ und „Push-to-Web“

Weiterhin wurde untersucht, ob die Teilnehmerinnen und Teilnehmer in den Gruppen „Mixed-Mode“ und „Push-to-Web“, die den Fragebogen entweder postalisch oder online ausgefüllt haben, sich hinsichtlich bestimmter Charakteristika unterscheiden. Aufgrund der relativ geringen Fallzahlen hinsichtlich der Beantwortung dieser spezifischen Fragestellung, wurden keine multiplen Regressionsmodelle gerechnet. Stattdessen wurden nur Haupteffekte analysiert.

Tabelle 15 zeigt, ob sich die Teilnehmerinnen und Teilnehmer, die den Fragebogen postalisch ausgefüllt haben, bezüglich ihres Alters, ihres Geschlechts²⁸, ihres Bildungsabschlusses oder hinsichtlich der Anzahl ihrer Komorbiditäten von den Personen, die sich für die Onlineversion entschieden haben, unterscheiden. Für die Variablen Alter und Anzahl der Komorbiditäten wurden zweiseitige t-Tests gerechnet und für die Variablen Geschlecht und Bildung wurden aufgrund geringer Zellenausprägung zweiseitige exakte Tests nach Fisher gerechnet.

Es zeigt sich, dass in der Gruppe „Mixed-Mode“ ältere Personen sowie Personen mit niedrigerem Bildungsabschluss die postalische Fragebogenversion präferierten. In der „Push-to-Web“-Gruppe zeigt sich ebenso, dass ältere Personen die Befragung eher postalisch beantworteten als online, jedoch gab es keinen signifikanten Unterschied. Bei den Variablen Geschlecht und Anzahl der Komorbiditäten konnten weder bei „Mixed-Mode“ noch bei „Push-to-Web“ signifikante Unterschiede festgestellt werden. Somit können die Ergebnisse aus Kapitel 6 teilweise bestätigt werden. Alter und Bildung scheinen einen Einfluss auf die Internetnutzung bzw. auf die Wahl zwischen Onlinefragebogen und Papierfragebogen zu haben. Dass Geschlecht und die Anzahl der Komorbiditäten keinen signifikanten Einfluss bei der vorliegenden Analyse haben, ist möglicherweise auf die geringe Stichprobengröße zurückzuführen.

Tabelle 15: Moduspräferenzen in den Studienarmen „Mixed-Mode“ und „Push-to-Web“

Charakteristika	Mixed-Mode			Push-to-Web		
	Papier	Online	P-Wert	Papier	Online	P-Wert
Alter			p < 0,000			p = 0,008
MW (SD)	69,8 (12,3)	55 (11,8)		69,9 (12,4)	60,7 (16)	
Geschlecht			p = 0,348			p = 1
Weiblich; n (%)	29 (40,8)	7 (53,8)		12 (44,4)	19 (42,2)	
Männlich; n (%)	42 (59,2)	5 (38,5)		15 (55,6)	25 (55,6)	
Divers; n (%)	0 (0)	1 (7,7)		0 (0)	1 (2,2)	
Höchster schulischer Bildungsabschluss*			p = 0,003			p = 0,109
Niedriger schulischer Abschluss; n (%)	38 (60,3)	1 (8,3)		13 (61,9)	16 (36,4)	

²⁸ Da insgesamt nur zwei Personen ihr Geschlecht als „divers“ angaben, konnten diese beiden Fälle bei der Analyse von Unterschieden nach Geschlecht nicht berücksichtigt werden.

Charakteristika	Mixed-Mode			Push-to-Web		
	Papier	Online	P-Wert	Papier	Online	P-Wert
Mittlerer schulischer Abschluss; n (%)	16 (25,4)	6 (50)		4 (19)	19 (43,2)	
Hoher schulischer Abschluss; n (%)	9 (14,3)	5 (41,7)		4 (19)	9 (20,5)	
Komorbiditäten			p = 0,07			p = 0,163
MW (SD)	3,5 (2)	2,7 (1,2)		2,5 (1,6)	3,1 (1,7)	
Gesamt; n (%)	73 (85 %)	13 (15 %)		28 (38 %)	45 (62 %)	

Abkürzungen: Med. = Median; MW = Mittelwert; SD = Standardabweichung; n = Anzahl

* Schulischer Bildungsabschluss wurde in drei Kategorien aufgeteilt: *niedriger schulischer Abschluss*: von der Schule abgegangen ohne Abschluss / Haupt- oder Volksschulabschluss / Abschluss der Polytechnischen Oberschule der DDR, 8. oder 9. Klasse / noch Schülerin bzw. Schüler einer allgemeinbildenden Vollzeitschule; *mittlerer schulischer Abschluss*: Realschulabschluss/Abschluss der Polytechnischen Oberschule der DDR, 10. Klasse; *hoher schulischer Abschluss*: Fachhochschulreife, den Abschluss einer Fachoberschule/allgemeine oder fachgebundene Hochschulreife / Abitur (Gymnasium bzw. EOS, auch EOS mit Lehre).

9.3 Item-Nonresponse

Ein Ziel der vorliegenden Studie war es, zu untersuchen, ob sich die Datenqualität, definiert als der Anteil der Item-Nonresponse („Weiß nicht mehr“-Antworten und Missings) zwischen dem Onlinefragebogen und dem Papierfragebogen unterscheidet. Um zu untersuchen, ob der Modus einen signifikanten Einfluss auf den Anteil an Nonresponse hat bzw. ob sich dieser etwaige Unterschied durch Kontrollvariablen erklären lässt, wurden quasibinomial-logistische Regressionsmodelle gerechnet.²⁹ Als abhängige Variablen wurden zum einen die Anzahl an „Weiß nicht mehr“-Antworten im Fragebogen in Relation zur maximal möglichen Anzahl an „Weiß nicht mehr“-Antworten genommen und zum anderen die Anzahl von Missings in Relation zur maximalen Anzahl an Fragen im Fragebogen. Da in der Onlineversion des Fragebogens die Fragen mit der Sprunganweisung ausgefüllt werden mussten, weil der Onlinefragebogen sonst nicht zu korrekten Fragen hätte weiterspringen können, wurde das Modell ohne die Fragen mit der Sprunganweisung gerechnet, damit sowohl in der Onlineversion als auch in der Papierversion die maximal mögliche Anzahl von Missings identisch ist. Als Kontrollvariablen wurden wie in Kapitel 6 das Alter, das Geschlecht, die Bildungsniveau, sowie die Anzahl der Komorbiditäten herangezogen.

²⁹ Aufgrund von Überdispersion der Daten wurden quasibinomial-logistische Regressionsmodelle gerechnet.

Bezüglich des Einflusses des Modus auf die Wahrscheinlichkeit für eine „Weiß nicht mehr“-Antwort konnte kein signifikanter Effekt des Erhebungsmodus festgestellt werden (siehe Abbildung A 2 im Anhang). Bezüglich der Wahrscheinlichkeit, eine Frage gar nicht zu beantworten, wurde hingegen ein signifikanter Effekt des Modus aufgezeigt. So ist die Wahrscheinlichkeit, ein Item nicht zu beantworten, deutlich höher, wenn der Fragebogen in Papierform ausgefüllt wurde, im Vergleich zur Onlineversion (Abbildung 16). Daraus lässt sich schließen, dass es im Onlinefragebogen unwahrscheinlicher ist, eine Frage zu überspringen als im Papierfragebogen, und damit die Datenqualität in Bezug auf Item-Nonresponse erhöht wird. Dies könnte daran liegen, dass im Onlinefragebogen die Fragen hauptsächlich einzeln präsentiert werden und somit die Wahrscheinlichkeit niedriger ist, Fragen zu übersehen, da die Fragen stärker im Fokus stehen.

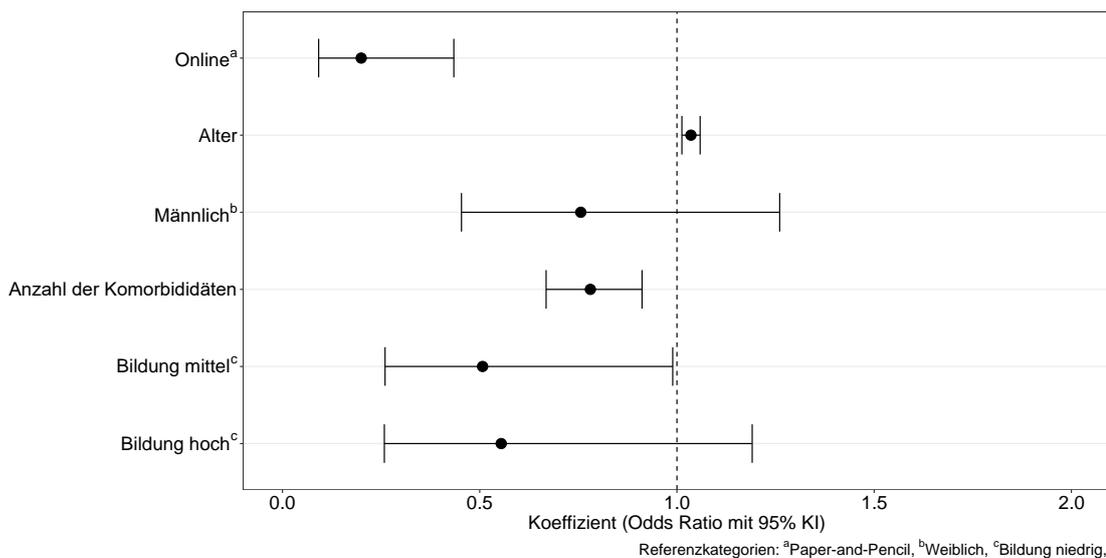


Abbildung 16: Modell zum Zusammenhang zwischen dem Erhebungsmodus und der Anzahl von Missings unter Kontrolle von Alter, Geschlecht, Anzahl der Komorbiditäten und Bildung

9.4 Prüfung systematischer Unterschiede bei der Beantwortung der Fragebögen

Zur Feststellung der Übereinstimmung der Befragungsergebnisse zwischen den Online- und Papierfragebögen wurden quasibinomial-logistische Regressionsmodelle berechnet.³⁰ Dabei wurden die zu den 13 Qualitätsindikatoren gehörenden Personen-Scores jeweils als abhängige Variable in separate Regressionsmodelle aufgenommen.³¹ Bei der Modellierung wurde schrittweise vorgegangen. Im ersten Schritt wurde als unabhängige Variable ausschließlich der Erhebungsmodus (online versus papierbasiert) ins Modell aufgenommen (Modell 1). Da der Modus nicht zufällig zugeteilt wurde und von verschiedenen Faktoren abhängig sein kann (siehe Abschnitt 9.2.3), wurden jeweils in einem zweiten Modell Kontrollvariablen (Geschlecht, Alter, Komorbiditäten, Bildungsstand) inkludiert (Modell 2), die sowohl mit dem Erhebungsmodus als auch mit dem Antwortverhalten assoziiert sein könnten (siehe Abschnitt 6.5). In Abbildung 17 sind die Koeffizienten des Modus aus den jeweiligen schrittweisen Modellen dargestellt. Es zeigt

³⁰ Aufgrund von Überdispersion der Daten wurden quasibinomial-logistische Regressionsmodelle gerechnet.

³¹ Die Hintergründe zur Berechnung der Personen-Scores sind in Abschnitt 5.8 beschrieben.

sich in beiden Modellen, dass der Modus keinen signifikanten Effekt auf die Ergebnisse der Befragung zu haben scheint. Alle Konfidenzintervalle umschließen, mit und ohne Berücksichtigung der Kontrollvariablen, den Odds Ratio-Wert von 1. Weiterhin liegen die Punktschätzer vergleichsweise nahe an 1. Daraus lässt sich schließen, dass die Wahrscheinlichkeit, dass der Modus einen systematischen Einfluss auf die Ergebnisse hinsichtlich der befragungsbasierten Qualitätsindikatoren hat, sehr gering ist.³²

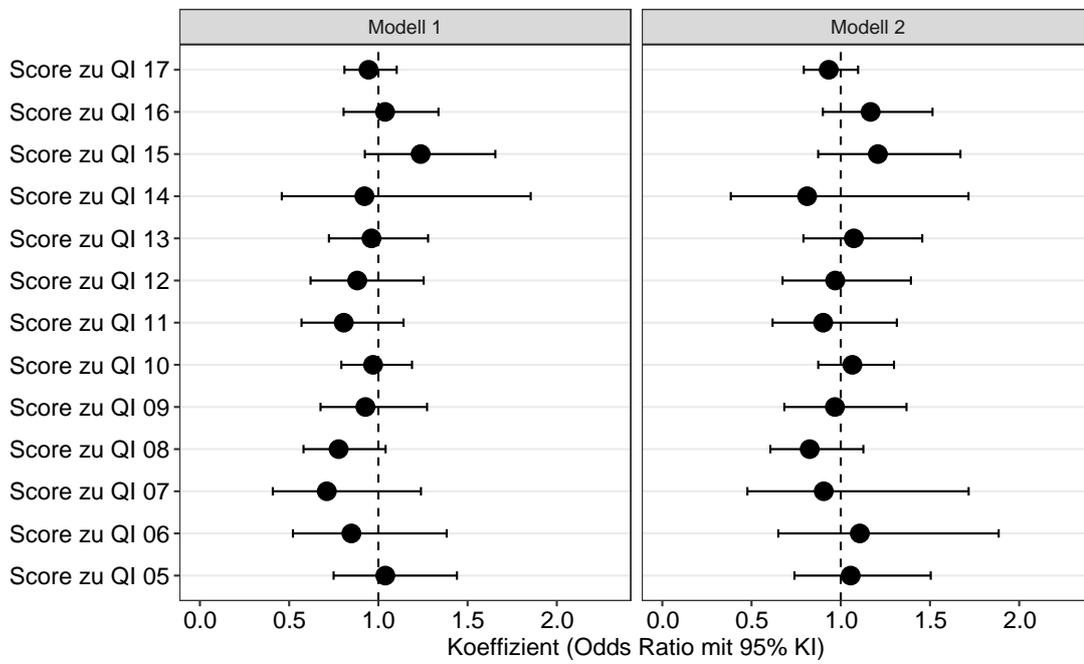


Abbildung 17: Einfluss des Modus auf die Personen-Scores, die auf Basis der zugrunde liegenden Items von 13 Qualitätsindikatoren der Patientenbefragung QS NET (Bereich Dialyse) berechnet wurden

9.5 Zusammenfassung

Die empirische Testung wurde durchgeführt, um zu überprüfen, inwieweit der Befragungsmodus bzw. die Einladungsform mit der Rücklaufquote, der Zusammensetzung der Nettostichproben sowie dem Antwortverhalten der Teilnehmerinnen und Teilnehmer assoziiert sind.

Die Rücklaufquoten waren für die Einladungsform „Single Mode Paper“ (Kontrollarm) am höchsten und für die Einladungsform „Single Mode Web“ am geringsten. Die Rücklaufquoten der experimentellen Arme „Push-to-Web“ und „Single Mode Web“ waren signifikant niedriger als die Rücklaufquoten im Kontrollarm. Die Rücklaufquote der experimentellen Bedingung „Mixed-Mode“ unterschied sich dagegen nicht signifikant von der Rücklaufquote der bisher verwendeten Einladungsform „Single Mode Paper“.

³² Als Sensitivitätsanalyse wurden alle Modelle einmal mit einer zusätzlichen Leistungserbringer-Variable als Fixed-Effect berechnet. Dabei zeigte sich, dass sich die Richtung und die Stärke der Effekte nur marginal von den Modellen ohne einen modellierten Leistungserbringer-Effekt unterscheiden.

Anhand der beiden Gruppen mit freier Moduswahl („Mixed-Mode“ und „Push-to-Web“) konnte gezeigt werden, dass die Befragten, die sich für den Online-Modus entschieden, signifikant jünger waren als jene, die sich für den Papierfragebogen entschieden.

Darüber hinaus war bei den Teilnehmerinnen und Teilnehmern, die den Fragebogen in Papierform ausgefüllt hatten, die Wahrscheinlichkeit höher, eine Frage nicht zu beantworten (Item-Nonresponse) als bei den Personen, die den Fragebogen online ausgefüllt hatten, was als höhere Datenqualität beim onlinebasierten Erhebungsmodus gewertet wurde.

Im Rahmen der Überprüfung des Einflusses des Modus auf die Personen-Scores, die auf Basis der zugrunde liegenden Items von 13 Qualitätsindikatoren der Patientenbefragung *QS NET* (Bereich Dialyse) berechnet wurden, konnte kein Effekt des Modus identifiziert werden.

Teil III: Empfehlungen zur Umsetzung von Onlinebefragungen im Rahmen der gesetzlichen Qualitätssicherung

10 Empfehlung zum Einsatz von onlinebasierten Patientenbefragungen im Regelbetrieb

10.1 Empfehlungen zum Befragungsmodus

Übergeordnetes Ziel des vorliegenden Projekts war gemäß Beauftragung durch den G-BA die Erarbeitung von Empfehlungen, wie Onlinebefragungen im Rahmen der sektorenübergreifenden Qualitätssicherung zukünftig umgesetzt werden sollen, um Patientenbefragungen effizienter zu gestalten. Als Voraussetzung dafür wurden auf Basis einer orientierenden Recherche zuerst die zur Verfügung stehenden Umsetzungsmöglichkeiten von Onlinebefragungen identifiziert und entsprechende Vor- und Nachteile gegenübergestellt (Zielsetzung I.a.; vgl. Abschnitt 3.2 sowie Kapitel 7). Diesbezüglich kann grundsätzlich zwischen Single-Mode-Onlinebefragungen und Mixed-Mode-Befragungen unterschieden werden. Eine Onlinebefragung im Single-Mode würde dabei die papierbasierte Patientenbefragung, wie sie aktuell im Rahmen der gesetzlichen QS vorgesehen ist, gänzlich ersetzen. Bei Mixed-Mode-Befragungen bliebe die papierbasierte Befragung bestehen und würde um die Möglichkeit der Beantwortung des jeweiligen Fragebogens im Onlineformat ergänzt werden, je nach individueller Präferenz bzw. Vorliegen eines Internetzugangs sowie anderer notwendiger Voraussetzungen hinsichtlich Medienverfügbarkeit und entsprechender Kompetenzen der Befragten.

10.1.1 Bewertung der Umsetzbarkeit und der rechtlichen Rahmenbedingungen unterschiedlicher Befragungsmodi

Unter der Voraussetzung, dass die Kontaktaufnahme der Patientinnen und Patienten weiterhin per postalischem Anschreiben erfolgt (siehe hierzu Abschnitt 10.2), wurden Onlinebefragungen als Instrumente der gesetzlichen QS im Regelbetrieb grundsätzlich sowohl im Single-Mode als auch im Mixed-Mode als umsetzbar bewertet (Zielsetzung I.b.; vgl. Abschnitte 7.3.3 und 7.3.4). Beide Szenarien erfordern in diesem Zusammenhang geringfügige Anpassungen bei den Prozessen der Versendestelle der Patientenbefragung, die voraussichtlich relativ einfach zu implementieren sind. Im Gegensatz dazu ist aufgrund der Umstellung auf onlinebasierte Patientenbefragungen allerdings mit erheblichen Änderungen hinsichtlich der Aufgaben der Fragebogenannahmestelle zu rechnen, was ebenfalls sowohl für Onlinebefragungen im Single-Mode als auch für Mixed-Mode-Befragungen gilt. Um die identifizierten IT-technischen Anforderungen erfüllen zu können (Zielsetzung I.c.; vgl. Abschnitt 7.3.5), müsste aus Perspektive des IQTIG aus diesem Grund in beiden Fällen ggf. eine separate Fragebogenannahmestelle für onlinebasierte Patientenbefragungen eingerichtet bzw. das bisherige Aufgabenspektrum der Fragebogenannahmestelle deutlich erweitert werden.

Hinsichtlich der Prüfung der gesetzlichen und datenschutzrechtlichen Rahmenbedingungen von Onlinebefragungen im Kontext der gesetzlichen Qualitätssicherung (Zielsetzung I.d.) konnten für beide Umsetzungsmöglichkeiten keine grundsätzlichen Hindernisse identifiziert werden, die der Etablierung von onlinebasierten Befragungen entgegenstehen, was insbesondere für die posta-

liche Kontaktaufnahme der Befragten gilt (vgl. Abschnitt 7.4). Bei elektronischer Kontaktaufnahme hingegen wären personenbezogenen Daten erforderlich (z. B. E-Mail-Adressen oder Telefonnummern von Patientinnen und Patienten), die gesondert erhoben werden müssten, wobei hierzu ggf. erst eine Einwilligung eingeholt werden müsste (siehe hierzu auch Abschnitt 10.2).

Zusammengefasst kann auf Basis der Bewertung der Umsetzbarkeit und der gesetzlichen und datenschutzrechtlichen Rahmenbedingungen keine eindeutige Präferenz für eine der beiden grundlegenden Formen von Onlinebefragungen identifiziert werden.

Sowohl Single-Mode-Onlinebefragungen als auch Mixed-Mode-Befragungen scheinen im Rahmen der gesetzlichen QS grundsätzlich umsetzbar.

10.1.2 Empfehlung einer Mixed-Mode-Onlinebefragung auf Basis der zu erwartenden Effekte und Konsequenzen für die Datenerhebung

Über die Bewertung der Umsetzbarkeit und der rechtlichen Rahmenbedingungen hinaus sollte im Rahmen des vorliegenden Projekts laut Beauftragung des G-BA eine Analyse von Effekten hinsichtlich Zielgruppenerreichung (Zielsetzung III.a.) und Antwortverhalten (Zielsetzung III.b.), die mit der Umsetzung als Onlinebefragung verbunden sind, erfolgen. Auf Basis entsprechender Ergebnisse sollte zudem eine Einschätzung dahingehend gegeben werden, ob und ggf. wie die Ergebnisse der onlinebasierten Patientenbefragung adjustiert werden sollen, um Vergleichbarkeit mit Paper-Pencil-Befragungen zu gewährleisten (Zielsetzung III.b.). Zur Beantwortung der damit verbundenen Fragestellungen wurden öffentlich zugängliche Daten zur Internetnutzung in Deutschland (siehe Kapitel 6), orientierende und systematische Literaturrecherchen (siehe Abschnitt 7.1 und Kapitel 8) sowie eine eigene empirische Untersuchung auf Basis von Primärdaten am Beispiel der Patientenbefragung *QS NET* (siehe Kapitel 9) herangezogen. Im Rahmen der Analysen zeigten sich erhebliche Unterschiede zwischen den verschiedenen Befragungsmodi hinsichtlich der Zielgruppenerreichung bzw. der zu erwartenden Rücklaufquoten zugunsten des Mixed-Mode-Ansatzes. Im Gegensatz dazu fanden sich keine Hinweise, dass der Befragungsmodus das Antwortverhalten auf eine Weise systematisch beeinflussen könnte, dass die Vergleichbarkeit der Ergebnisse der Patientenbefragung davon beeinträchtigt wäre.

Auf dieser Grundlage empfiehlt das IQTIG einen Mixed-Mode-Ansatz als präferierte Möglichkeit zur Durchführung von onlinebasierten Patientenbefragungen im Rahmen der gesetzlichen QS, bei der die Befragten selbst entscheiden können, ob sie die Befragung papierbasiert oder als Onlinefragebogen beantworten.

Diese Empfehlung soll im Folgenden auf Basis konkreter Ergebnisse des vorliegenden Projekts im Detail begründet werden:

- Die Analysen der Daten zur Internetnutzung in Deutschland (vgl. Kapitel 6) lassen den Schluss zu, dass mit einer Onlinebefragung im Single-Mode bestimmte Teile der Zielpopulation systematisch seltener erreicht werden könnten. Insbesondere bei Personen über 60 Jahren wurde ein Rückgang der Internetnutzung deutlich, der sich mit steigendem Alter weiter verstärkt. Bezogen auf die deutsche Gesamtbevölkerung zeigte sich für die Jahre 2018 und 2019, dass

in der Altersgruppe der Über-70-Jährigen nur etwa jede bzw. jeder zweite das Internet nutzte; in der Gruppe der Über-80-Jährigen war es nur noch jede bzw. jeder vierte. Auch wenn die Nutzung des Internets im Zeitverlauf innerhalb der letzten Jahre in allen Altersgruppen kontinuierlich zugenommen hat, ist bislang nicht abzusehen, bis wann alle Altersgruppen eine vergleichbar hohe Wahrscheinlichkeit aufweisen, das Internet zu nutzen. Darüber hinaus war die Internetnutzung unter statistischer Kontrolle des Alters negativ mit niedrigem Bildungsstand, weiblichem Geschlecht, der Anzahl an chronischen Erkrankungen sowie mit funktionellen Einschränkungen im Alltag assoziiert. Dies verdeutlicht, dass hinsichtlich der Repräsentativität von onlinebasierten Befragungen auch unabhängig vom Alter mit Coverage-Problemen zu rechnen ist. Dies dürfte in besonderem Maße für Befragungen von Patientinnen und Patienten im Rahmen der gesetzlichen QS gelten, da nicht nur die Altersverteilung der Zielpopulation in vielen Versorgungsbereichen im Vergleich mit der Allgemeinbevölkerung deutlich höher ausfällt, sondern in klinischen Populationen funktionelle Einschränkungen sowie eine hohe Anzahl an Komorbiditäten besonders häufig vorkommen. Zusammengefasst kann bei einer Single-Mode-Onlinebefragung nicht davon ausgegangen werden, dass alle zur Zielpopulation gehörenden Patientinnen und Patienten die gleiche Wahrscheinlichkeit aufweisen, an der Befragung teilnehmen zu können, was eine Grundvoraussetzung zur repräsentativen Abbildung der patientenberichteten Behandlungsqualität darstellt. Bei Mixed-Mode-Befragungen ist diese Problematik durch die zusätzliche Möglichkeit zur Beantwortung eines Papierfragebogens nicht zu erwarten.

- Zusätzlich zur Coverage-Problematik zeigen die Analysen der Daten zur Internetnutzung in Deutschland (vgl. Kapitel 6) auf, dass insgesamt schlichtweg weniger Patientinnen und Patienten die Möglichkeit haben, an einer onlinebasierten Patientenbefragung teilzunehmen, als dies auf Basis einer reinen Paper-Pencil-Befragung oder einer Mixed-Mode-Befragung der Fall ist.
- Sowohl die Ergebnisse der orientierenden Recherche (vgl. Abschnitt 7.1) als auch die Ergebnisse der Übersicht von systematischen Reviews (vgl. Abschnitt 8.3.1) deuten vereinzelt darauf hin, dass insbesondere sehr junge, nicht klinische Bevölkerungsgruppen ggf. besser mit Onlinebefragungen im Single-Mode zu erreichen sein können, als mit reinen postalischen Paper-Pencil-Befragungen. Demgegenüber finden sich in den Ergebnissen aus der Übersicht von systematischen Reviews für gesundheitsbezogene Befragungen Hinweise dafür, dass in der Allgemeinbevölkerung bzw. in klinischen Populationen elektronische bzw. onlinebasierte Befragungen häufig mit niedrigeren Rücklaufquoten assoziiert sind, verglichen mit papierbasierten Befragungen. Dies stimmt mit dem Ergebnis aus der orientierenden Recherche überein, dass sich bei reinen Onlinebefragungen im Allgemeinen niedrigere Rücklaufquoten finden.
- Im Rahmen der orientierenden Recherche (vgl. Abschnitt 7.1) wurde deutlich, dass Vorteile von Onlinebefragungen im Single-Mode zumindest zum Teil auch auf Mixed-Mode-Befragungen übertragen werden können. Darüber hinaus werden Mixed-Mode-Befragungen als potenzielle Lösung für die oben beschriebene Coverage-Problematik bei reinen Onlinebefragungen sowie auch als Möglichkeit zur Erhöhung der Rücklaufquote diskutiert. Einschränkend hierzu zu erwähnen ist, dass die Evidenzlage nicht eindeutig ist und vom jeweiligen Studienkontext abzuhängen scheint.

- Im Hinblick auf diese Limitationen wurden im vorliegenden Projekt eigene empirische Untersuchungen durchgeführt (vgl. Kapitel 9), um möglichst ähnliche Bedingungen wie im Regelbetrieb der gesetzlichen QS zu schaffen. Dabei stellte sich heraus, dass die Rücklaufquote bei einer Onlinebefragung im Single-Mode um mehr als 30 Prozentpunkte – und damit statistisch signifikant – niedriger ausfiel, als die herkömmliche papierbasierte Befragung. Die Rücklaufquote der simultanen Mixed-Mode-Befragung unterschied sich nicht signifikant von der herkömmlichen papierbasierten Variante. Anzumerken ist, dass auch mit der sequenziellen Mixed-Mode-Befragung eine statistisch signifikant niedrigere Rücklaufquote assoziiert war (vgl. Abschnitt 9.2).
- Die Ergebnisse der empirischen Untersuchung (vgl. Abschnitt 9.2) deuteten insgesamt auf eine deutliche Präferenz der Befragten für die papierbasierte Fragebogenversion hin. So wählten in der Teilgruppe der Patientinnen und Patienten, die der simultanen Mixed-Mode-Befragung zugeordnet wurde, nur 15 % die Beantwortung im Onlineformat, während 85 % den Papierfragebogen postalisch beantworteten. Die Gruppe derer, die sich für das Onlineformat entschieden, war dabei deutlich jünger.
- Weder die Ergebnisse der Übersicht von systematischen Reviews (vgl. Abschnitt 8.3.1) noch die Ergebnisse der empirischen Untersuchung (vgl. Abschnitt 9.4) deuten darauf hin, dass mit systematischen Antwortverzerrungen zu rechnen ist, wenn unterschiedliche Befragungsmodi verwendet werden. Da somit keine Moduseffekte zu erwarten sind, scheint eine Mixed-Mode-Befragung zulässig, ohne die onlinegenerierten Daten statistisch adjustieren zu müssen.

Darüber hinaus wird hinsichtlich des Erstanschreibens und der Erinnerungsschreiben auf Basis der gewonnenen Erkenntnisse empfohlen, für Patientenbefragungen im Rahmen der externen QS in aller Regel einen *simultanen* Mixed-Mode-Ansatz zu verfolgen.

Begründet wird diese Empfehlung folgendermaßen:

- In der empirischen Untersuchung (vgl. Abschnitt 9.2) war mit dem sequenziellen Mixed-Mode-Ansatz („Push-to-Web“) eine um fast 20 Prozentpunkte niedrigere Rücklaufquote assoziiert, verglichen mit dem herkömmlichen papierbasierten Ansatz. Dieses Ergebnis war statistisch signifikant, was nicht auf die erreichte Rücklaufquote in dem simultanen Mixed-Mode-Ansatz zutraf. Zudem beantwortete in dem simultanen Mixed-Mode-Ansatz ein wesentlich größerer Anteil der Antwortenden den Fragebogen bereits im Zeitraum zwischen dem Erstanschreiben und dem ersten Erinnerungsschreiben (70 %), als dies in dem Push-to-Web-Ansatz (47 %) der Fall war. Dies ist u. a. deshalb relevant, da in diesen Fällen keine weiteren Kosten für den Druck und die Aussendung von Erinnerungsschreiben und weiteren Papierfragebögen entstehen.
- Auf Basis der Analysen der Daten zur Internetnutzung in Deutschland (vgl. Kapitel 6) und der orientierenden sowie der systematischen Recherchen (vgl. Abschnitte 7.1 und 8.3.1) ist davon auszugehen, dass insbesondere bei Befragungen von klinischen Populationen die Internetnutzung nicht in jedem Fall vorausgesetzt werden kann bzw. die Teilnahmebereitschaft an Onlinebefragungen weniger stark ausgeprägt ist. Damit übereinstimmend deuten auch die Ergebnisse der empirischen Untersuchung (vgl. Abschnitt 9.2) auf eine deutliche Präferenz der

Befragten für die papierbasierte Fragebogenversion hin. Es kann daher für die betroffenen Patientinnen und Patienten irritierend und ggf. auch belastend sein, im Zuge einer Push-to-Web-Befragung mehrmals die Aufforderung zur Teilnahme an einer Onlinebefragung zu erhalten, der sie aufgrund fehlender Voraussetzungen nicht nachkommen können.

Die Ergebnisse der orientierenden Recherche deuten darauf hin, dass für bestimmte Populationen eine sequenzielle Mixed-Mode-Befragung hinsichtlich der Teilnahmebereitschaft nicht in jedem Fall nachteilig gegenüber einem simultanen Ansatz sein muss (vgl. Abschnitt 7.1.2). Da sequenzielle Mixed-Mode-Befragungen insgesamt kostengünstiger ausfallen als simultane Mixed-Mode-Befragungen, kann für vereinzelte Zielpopulationen bzw. für bestimmte QS-Verfahren ein Push-to-Web-Ansatz zielführender sein als ein simultaner Mixed-Mode-Ansatz. Dies scheint vor allem bei Zielpopulationen mit verhältnismäßig jungen und wenig komorbiden Patientinnen und Patienten sinnvoll. Im Rahmen einer Zielgruppenanalyse für die jeweilige Entwicklung muss genauer geprüft werden, inwiefern sequenzielle Mixed-Mode-Befragungen vorteilhafter erscheinen (vgl. dazu auch Kapitel 11).

Für die Entwicklung von Patientenbefragungen im Rahmen von QS-Verfahren mit bestimmten Zielpopulationen (z. B. mit jüngeren, wenig komorbiden Patientinnen und Patienten) wird daher empfohlen, im Entwicklungsprozess zu prüfen, ob der Einsatz eines generellen Push-to-Web-Ansatzes oder aber eine Kombination aus beiden Mixed-Mode-Ansätzen (sog. „Targeting“, vgl. Abschnitt 7.1.2) zielführend ist.

10.2 Empfehlungen zur Kontaktierung der Befragten

Für die bisher entwickelten papierbasierten Patientenbefragungen wurde empfohlen, dass zum Zeitpunkt des Fragebogenversands Fragebögen, Anschreiben, vorfrankierte Rücksendeküverts und Informationsschreiben in gedruckter Form postalisch an die adressierten Patientinnen und Patienten gesendet werden sollen (IQTIG 2018b, IQTIG 2018a, IQTIG 2020). Zwei Wochen nach dem Versand der Fragebögen sollen jene Patientinnen und Patienten, von denen bis zu diesem Zeitpunkt kein Fragebogeneingang verzeichnet werden konnte, von der Versendestelle Patientenbefragung postalisch ein kontrolliertes Erinnerungsschreiben erhalten. Kontrolliert bedeutet, dass der Versand des ersten Erinnerungsschreibens an eine Prüfung des Fragebogeneingangs bei der Fragebogenannahmestelle des IQTIG gekoppelt ist. Zur Feststellung des Fragebogeneingangs übermittelt die Fragebogenannahmestelle kontinuierlich alle Fragebogen-IDs der eingegangenen Fragebögen an die Versendestelle. In weiteren zwei Wochen sollen Patientinnen und Patienten, von denen bis dahin noch kein Fragebogen in der Fragebogenannahmestelle des IQTIG eingegangen ist, ein zweites kontrolliertes Erinnerungsschreiben inklusive des kompletten Fragebogens sowie des Rücksendeküverts erhalten.

Auf Basis der im vorliegenden Projekt gewonnenen Erkenntnisse konnten keine Hinweise dafür gefunden werden, dass dieser zweiwöchige Kontaktierungs- bzw. Erinnerungsrhythmus, bestehend aus einem Erstanschreiben und zwei kontrollierten Erinnerungsschreiben, im Rahmen einer onlinebasierten Patientenbefragung weniger zielführend wäre, als im Rahmen der herkömmlichen papierbasierten Befragungen. Daher soll an diesem Ablauf auch mit der empfohlenen Mixed-Mode-Befragung (vgl. Abschnitt 10.1) weiterhin festgehalten werden.

Im Fall einer simultanen Mixed-Mode-Befragung sollen in Anlehnung an die bisherigen Empfehlungen zur papierbasierten Patientenbefragung zum Zeitpunkt des Erstanschreibens neben den Zugangsdaten zur optionalen Onlinebefragung weiterhin Papierfragebögen, Anschreiben, vorfrankiertes Rücksendekuvert und Informationsschreiben postalisch zugesendet werden. Der Papierfragebogen sowie das zugehörige Rücksendekuvert sollen im Zuge des zweiten Erinnerungsschreibens erneut versendet werden. Aufgrund des sehr geringen Zusatzaufwands sollen die Zugangsdaten für die Patientinnen und Patienten zur Onlinebefragung hingegen sowohl im Erstanschreiben als auch im Zuge beider Erinnerungsschreiben enthalten sein.

Falls eine sequenziellen Mixed-Mode-Befragung („Push-to-Web“; vgl. Abschnitt 10.1) durchgeführt werden soll, sollen die Patientinnen und Patienten zum Zeitpunkt des Erstanschreibens neben den Zugangsdaten zur Onlinebefragung nur ein Anschreiben und ein Informationsschreiben erhalten. Im Rahmen der beiden Erinnerungen sollen neben den Erinnerungsschreiben die Zugangsdaten zur Onlinebefragung erneut versendet werden. Zum Zeitpunkt des zweiten Erinnerungsschreibens sollen die Patientinnen und Patienten zusätzlich einen optionalen Papierfragebogen sowie ein zugehöriges vorfrankiertes Rücksendekuvert erhalten.

Darüber hinaus soll die Kontaktaufnahme zu den Befragten bei der Umsetzung einer Mixed-Mode-Befragung grundsätzlich postalisch erfolgen (siehe auch Abschnitt 7.3.2). Dies gilt sowohl für die Erstanschreiben als auch für die beiden Erinnerungsschreiben und soll auch im Fall eines Push-to-Web-Ansatzes umgesetzt werden.

Begründet wird diese Empfehlung auf Basis der folgenden Erkenntnisse aus dem vorliegenden Projekt:

- Da ein Papierfragebogen mitgeschickt werden muss, ist für Mixed-Mode-Befragungen in der Regel eine postalische Kontaktierung sinnvoll und stellt das übliche Vorgehen dar (vgl. Abschnitt 7.1.2).
- Wie bereits in Abschnitt 10.1.2 in Bezug auf Onlinebefragungen im Single-Mode dargestellt, zeigen die Daten zur Internetnutzung in Deutschland (vgl. Kapitel 6) auf, dass mit einer elektronischen Kontaktaufnahme bestimmte Teile der Zielpopulation schon allein deshalb systematisch seltener befragt werden würden, weil der Link zur Onlinebefragung nicht genutzt werden könnte. Darüber hinaus zeigte sich, dass es selbst unter den Internetnutzerinnen und Internetnutzern in allen Altersgruppen einen nicht zu vernachlässigenden Anteil an Personen gibt, die E-Mail-Dienste nur selten oder gar nicht nutzen und dass dieser Anteil mit zunehmendem Alter und niedrigem Bildungsniveau steigt. Ähnliches gilt für die Nutzung von Smartphones, die neben der Kontaktierung per E-Mail eine zweite Möglichkeit zur elektronischen Kontaktaufnahme darstellen. Auch hier zeigen die Daten auf, dass die Nutzung eines Smartphones selbst unter Internetnutzerinnen und Internetnutzern, insbesondere in höheren Altersgruppen, aber auch bei Jüngeren, keinesfalls vorausgesetzt werden kann. Diese Befunde verdeutlichen, dass bei elektronischer Kontaktaufnahme mit verstärkten Problemen hinsichtlich Coverage und kleinen Fallzahlen in einzelnen Einrichtungen zu rechnen ist.

- Die Nutzung der Adressdaten zur Kontaktaufnahme mit der Patientin bzw. dem Patienten ist bereits in den gesetzlichen Grundlagen verankert und datenschutzrechtlich zulässig. Zwar wäre nach § 299 Abs. 4 SGB V aus rechtlicher Perspektive grundsätzlich auch die Übermittlung von Telefonnummern oder E-Mail-Adressen von Versicherten möglich, allerdings handelt es sich bei solchen personenbezogenen Daten nicht um Versichertenstammdaten (vgl. Abschnitt 7.2). Es ist daher anzunehmen, dass solche Angaben in nur sehr geringem Umfang bei den Leistungserbringern vorliegen und diese daher gesondert erhoben und ggf. hierzu erst Einwilligungen eingeholt werden müssten. Dies könnte die Anzahl der Patientinnen und Patienten, die im Rahmen einer Patientenbefragung kontaktiert werden können, weiter reduzieren. Zudem wäre mit Erteilung der Einwilligung das Prinzip aufgelöst, dass jede Patientin bzw. jeder Patient des betreffenden QS-Verfahrens potenziell in die Bewertung der Behandlungsqualität eingehen kann.
- Die Adressdaten für die postalische Kontaktaufnahme können automatisiert von der eGK-Versichertenkarte der Patientin oder des Patienten ausgelesen und genutzt werden, wodurch das Fehlerrisiko minimiert wird. Im Gegensatz dazu würde die notwendige manuelle Erhebung von E-Mail-Adressen und/oder Telefonnummern für eine elektronische Kontaktierung ein erhöhtes Fehlerrisiko sowie einen hohen Zusatzaufwand seitens der Leistungserbringer bedeuten.
- Der Grundsatz der Datensparsamkeit erfordert, dass die Verarbeitung zusätzlicher Patientendaten – wie E-Mail-Adressen oder Telefonnummern – sehr gut begründet werden müsste. Demgegenüber stehen die Ergebnisse der wissenschaftlichen Literatur (vgl. Kapitel 7 und 8), dass postalische bzw. papierbasierte Befragungen und Mixed-Mode-Befragungen mit vielen Vorteilen, insbesondere hinsichtlich der Teilnahmebereitschaft bzw. Response-Raten, einhergehen.
- Einladungen zur Teilnahme an der Befragung auf elektronischem Wege können relativ leicht verloren gehen, z. B. durch SPAM-Filter der genutzten E-Mail-Anbieter, was die Anzahl der Teilnehmerinnen und Teilnehmer weiter reduzieren würde.
- Für Befragungen im Single-Mode wird häufig empfohlen, für Kontaktierung und Befragung den gleichen Modus zu nutzen, um einen Medienbruch zu vermeiden. Allerdings finden sich in der Literatur auch Hinweise, dass postalische Kontaktaufnahme bei Onlinebefragungen die Rücklaufquote gegenüber digitaler Kontaktaufnahme sogar erhöhen kann. Begründet wird dies nicht nur mit der Tatsache, dass E-Mails die Adressatin bzw. den Adressanten in der Regel seltener erreichen als postalische Anschreiben, sondern auch damit, dass per Post verschickte Einladungen die wahrgenommene Seriosität bzw. die subjektive Bedeutung der Befragung erhöhen (vgl. Abschnitt 7.1.1).

10.3 Darstellung des prozessualen Ablaufs bei der Umsetzung einer Mixed-Mode-Befragung mit postalischer Kontaktierung

Eine mögliche technisch-organisatorische Realisierung einer Mixed-Mode-Befragung mit postalischer Kontaktierung ist in Abbildung 18 dargestellt. Die Abbildung enthält eine vereinfachte Darstellung der Datenflüsse, in denen die Details zur Verschlüsselung der Daten zugunsten der

Lesbarkeit und Übersichtlichkeit ausgespart wurden. Die entsprechenden Angaben zu den aktuellen Vorgaben hinsichtlich des Datenflusses bei Patientenbefragungen gemäß DeQS-RL sind in Abschnitt 7.3.1 dargestellt (vgl. Abbildung 14) und anschließend detailliert erläutert.

Im Rahmen einer Mixed-Mode-Befragung mit postalischer Kontaktierung erfolgt die QS-Auslösung und die Zusammenstellung der für die Befragung relevanten Informationen beim Leistungserbringer identisch zur etablierten, papierbasierten Befragung. Auch der Versand der Daten an die Datenannahmestelle auf Landesebene, die Pseudonymisierung der Leistungserbringer und der Weiterversand der Daten an die Versendestelle Patientenbefragung sind identisch zum etablierten Verfahren.

Die Prozesse in der Versendestelle Patientenbefragung bezüglich Datenentgegennahme, Entschlüsselung, Datenprüfung und Stichprobenziehung folgen auch dem etablierten Vorgehen gemäß DeQS-RL. Die medizinischen Daten werden nach Dateneingang entschlüsselt, geprüft und zur Stichprobenziehung verwendet. Die Adressdaten werden im organisatorisch unabhängigen Bereich verschlüsselt gespeichert, bis die Patientinnen und Patienten in der Stichprobenziehung in das QS-Verfahren eingeschlossen wurden. Nach der Stichprobenziehung entschlüsselt die Versendestelle Patientenbefragung die Adressdaten zur Erstellung des Anschreibens.

Zur Durchführung einer Onlinebefragung benötigt die Versendestelle Patientenbefragung zusätzlich zu den bisher genutzten Informationen die Zugangsdaten zum Webportal, auf dem die Onlinebefragung durchgeführt wird. Das Webportal wird von der Fragebogenannahmestelle betrieben. Wie eine mögliche technische Umsetzung der Fragebogenannahmestelle für die Onlinebefragung erfolgen könnte, kann zum jetzigen Zeitpunkt noch nicht final festgelegt werden. Die Verantwortung für die Onlinebefragung liegt wie auch die der Annahme der papierbasierten Fragebögen beim IQTIG, unabhängig von der technischen Umsetzung. Folgende Möglichkeiten sind denkbar:

- Die Fragebogenannahmestelle für die papierbasierte Befragung ist in der Lage, nach Erweiterung des Auftragsgegenstands bzw. im Rahmen einer erweiterten Ausschreibung auch die Onlinebefragung durchzuführen.
- Ein weiterer Dienstleister übernimmt nach einer Ausschreibung die Funktion der Fragebogenannahmestelle für Onlinebefragungen.

Die Bereitstellung der Zugangsdaten für die Patientinnen und Patienten wird in Abbildung 18 durch den gelben Doppelpfeil symbolisiert. Sie kann auf unterschiedliche Art- und Weise erfolgen. Eine Möglichkeit ist die blockweise Erstellung von Zugangsdaten durch die Fragebogenannahmestelle (bspw. Erstellung von 10.000 Zugangsdaten auf einmal). Eine weitere Möglichkeit stellt die Bereitstellung der Zugangsdaten nach Bedarf dar. Bei dieser Möglichkeit fragt die Versendestelle die Fragebogenannahmestelle bei jedem Druck des Anschreibens einzeln an, sodass die Zugangsdaten für jeden Brief separat von der Fragebogenannahmestelle erstellt und an die Versendestelle Patientenbefragung übermittelt werden.

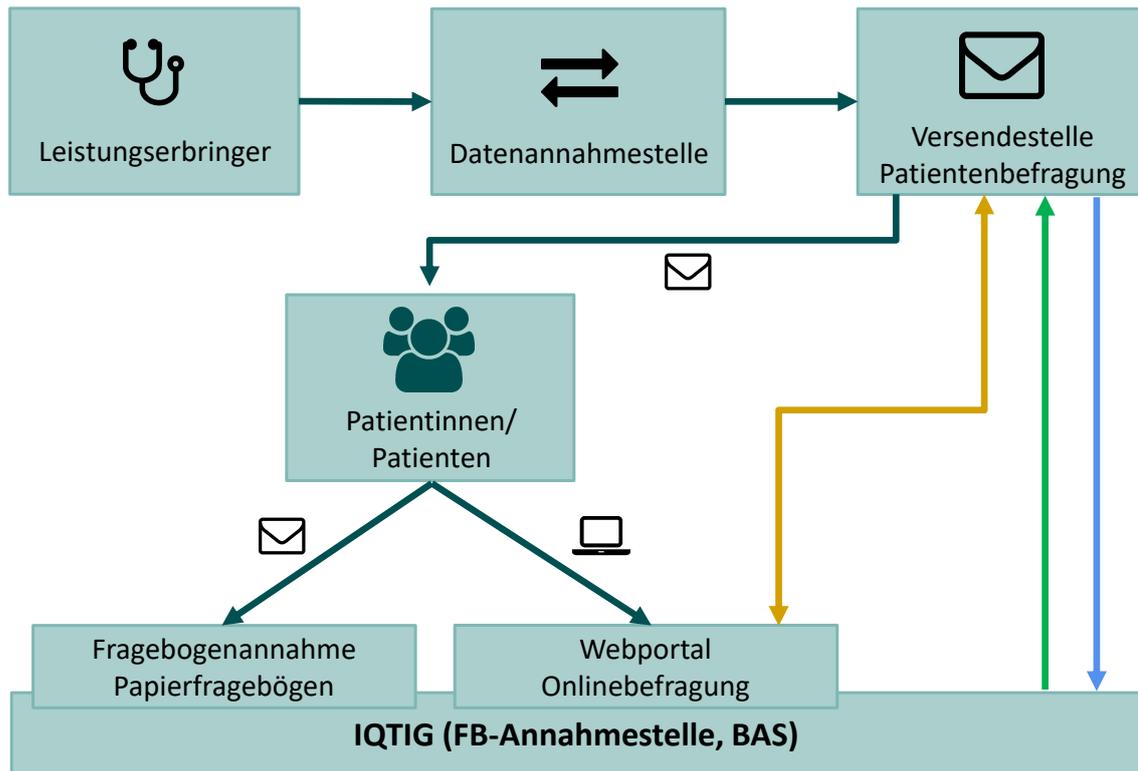


Abbildung 18: Prozessualer Ablauf bei der Umsetzung einer Mixed-Mode-Befragung mit postalischer Kontaktierung

Die Versendestelle Patientenbefragung erstellt das Anschreiben inklusive der Fragebogen-ID. Zusätzlich werden die Zugangsdaten zur Onlinebefragung auf das Anschreiben gedruckt. Die Zugangsdaten können als normaler Text, bspw. in der Form <http://fragebogenannahmestelle.de/pci/...> sowie als QR-Code auf das Anschreiben gedruckt werden. Patientinnen und Patienten mit einem Smartphone oder Tablet können mittels QR-Code die Befragungswebseite bequem öffnen. Je nach genutzter Technologie bezüglich der Zugangsdaten kann die Patientin oder der Patient über den Link direkt ins System eingeloggt werden, sodass Nutzernamen und Zugangspasswörter nicht vom Anschreiben abgetippt werden müssen.

Die Fragebogen-IDs der bereits beantworteten Fragebögen werden wie bei der papierbasierten Befragung täglich an die Versendestelle Patientenbefragung übermittelt, sodass Erinnerungsschreiben bzw. der Versand weiterer Papierfragebögen unterdrückt werden können (Abbildung 18, grüner Pfeil). Der Versand der Mapping-Tabellen, in denen die Fragebogen-ID, das Leistungserbringerpseudonym sowie ggf. behandlungsspezifische Daten von der Versendestelle Patientenbefragung an das IQTIG geschickt werden, erfolgt nach etabliertem Vorgehen (Abbildung 18, blauer Pfeil). Die Onlinebefragung hat keinen Einfluss auf diesen Prozessschritt.

Abschließend werden die Daten der papierbasierten und der onlinebasierten Befragung im IQTIG zusammengeführt und für die Auswertungen genutzt. Im Falle einer Übermittlung sowohl des Papierfragebogens als auch die Beantwortung der Befragung im Web ist zu entscheiden, welcher Fragebogen in die Auswertung miteinbezogen werden soll.

10.3.1 Vorteile und Einsparpotenziale der empfohlenen Umsetzung von onlinebasierten Patientenbefragungen

Der Einsatz von Mixed-Mode-Befragungen mit postalischer Kontaktierung für eine Patientenbefragung im Rahmen der gesetzlichen QS würde mit verschiedenen Vorteilen gegenüber der herkömmlichen papierbasierten Befragung verbunden sein. Diese Vorteile sollen im Folgenden kurz skizziert werden:

- Da im Rahmen von Mixed-Mode-Befragungen ein Teil der Patientinnen und Patienten den Fragebogen online beantwortet, ist mit deutlichen Einsparungen hinsichtlich der Kosten für die Dateneingabe zu rechnen, die nur für Paper-Pencil-Befragungen notwendig ist.
- Ob ein Fragebogen bereits beantwortet wurde, kann für Patientinnen und Patienten, die sich für das Onlineformat entscheiden, in Echtzeit festgestellt werden. Dies würde zu Vorteilen bei der Fragebogenlogistik und in weiterer Folge zu einer Kostenreduktion führen. Insbesondere würde durch die unmittelbare Datenübermittlung vermieden werden, dass Erinnerungsschreiben (ggf. mit Papierfragebogen) verschickt werden, obwohl sich der beantwortete Fragebogen bereits auf dem Postweg zur Fragebogenannahmestelle befindet.
- In der empirischen Untersuchung war die Rücklaufquote bei dem simultanen Mixed-Mode-Ansatz vor dem Zeitpunkt des ersten Erinnerungsschreibens um rund 10 Prozentpunkte höher als beim herkömmlichen Paper-Pencil-Ansatz. Dies lässt für den Regelbetrieb vermuten, dass Druck- und Portokosten für einen Teil der Erinnerungsschreiben eingespart werden könnten.
- Bei der Entscheidung für einen Push-to-Web-Ansatz würde es hingegen bereits beim ersten Anschreiben zu einer deutlichen Reduktion der Druckkosten kommen, da in diesem Fall der Fragebogen erst mit der zweiten Erinnerung postalisch verschickt werden soll.
- Durch die Vermeidung des Verschickens von Erinnerungsschreiben an Patientinnen und Patienten, die den Fragebogen bereits zurückgeschickt haben, könnten zudem Irritationen aufseiten der Betroffenen vermieden werden. In den Standard-Pretests der bisherigen Entwicklungsprojekte von Patientenbefragungen des IQTIG zeigte sich in diesem Zusammenhang regelmäßig, dass der Erhalt von Erinnerungsschreiben, obwohl der Fragebogen bereits zurückgeschickt wurde, der mit Abstand häufigste Grund war, weshalb Patientinnen und Patienten die für Rückfragen eingerichtete Hotline in Anspruch nahmen.
- Das Vermeiden der Übertragung der Daten eines Papierfragebogens in ein elektronisches Format führt zu einer Reduktion der Fehleranfälligkeit und erhöht die Datenqualität.
- Im Rahmen der empirischen Untersuchung zeigte sich, dass die Verwendung des Onlineformats zudem die Datenqualität in Bezug auf Item-Nonresponse erhöhen kann.
- Die Datenqualität kann bei onlinebasierten Befragungen zudem erhöht werden, indem komplexe Filterführungen in einfacher Weise umgesetzt werden können und die Eingabe unzulässiger bzw. unplausibler Werte verhindert werden kann.
- Bestimmte Zielgruppen können durch onlinebasierte Befragungen besser erreicht werden.

11 Konsequenzen für die Entwicklung und Validierung von onlinebasierten Befragungen

11.1 Konsequenzen für die Neuentwicklung von Befragungen im Rahmen der gesetzlichen Qualitätssicherung

Die allgemeinen Entwicklungsschritte von Patientenbefragungen als Instrumente der gesetzlichen QS sind in den Methodischen Grundlagen des IQTIG beschrieben (IQTIG 2019). Den Ausgangspunkt für die Entwicklung der Befragung bildet demnach regelmäßig ein versorgungsspezifisches Qualitätsmodell, für dessen Entwicklung Informationen aus einer Literaturrecherche sowie aus Fokusgruppen und ggf. leitfadengestützten Interviews herangezogen werden. Mit dem Qualitätsmodell werden verfahrensspezifische und für die Versorgungsqualität relevante Themen, sogenannte Qualitätsaspekte, identifiziert. Diese Qualitätsaspekte werden im nächsten Schritt durch sogenannte Qualitätsmerkmale konkretisiert bzw. ausdifferenziert. Die Qualitätsmerkmale müssen gemäß den Methodischen Grundlagen des IQTIG bestimmte Eignungskriterien erfüllen und stellen im Kontext der Patientenbefragung die Basis für die Generierung von Items dar, die in ihrer Gesamtheit den Fragebogen bilden (IQTIG 2019). Ein Expertengremium hat die Aufgabe, das IQTIG bei der Definition der Qualitätsmerkmale zu beraten. Auf Grundlage der Beratung durch das Expertengremium können die Qualitätsmerkmale ggf. modifiziert und abschließend definiert werden. Nach der Definition der Qualitätsmerkmale ist darüber hinaus ein schriftliches Stellungnahmeverfahren im Rahmen der Beteiligung nach § 137a Abs. 7 SGB V vorgesehen. Berechtigt sind hierfür die nach § 137a Abs. 7 SGB V zu beteiligenden Organisationen. Auf diese Weise können mögliche Anregungen der Beteiligungsberechtigten in die weiteren Arbeiten vor Abschluss der Entwicklung einfließen. Bis hierhin ergeben sich durch die Umstellung auf eine onlinebasierte Patientenbefragung bzw. auf eine Mixed-Mode-Befragung keine Änderungen im Entwicklungsprozess.

Nach dem Beteiligungsverfahren gemäß § 137a Abs. 7 SGB V beginnt die Entwicklung von Items auf Basis der Qualitätsmerkmale. Diese Items werden schließlich in Form eines Fragebogens zusammengestellt. Im Falle der Entwicklung von Mixed-Mode-Befragungen müssen neben dem Layout der Papierversion der Fragebögen auch Layout und Design der Onlineversion festgelegt und die Onlinebefragung entsprechend programmiert werden. Dabei sollten die in Abschnitt 7.2 vorgestellten Designmerkmale von Onlinefragebögen bzw. die spezifischen Hinweise in Bezug auf Mixed-Mode-Befragungen beachtet werden. Der Fragebogen wird danach in einer zweistufigen Pretestung, bestehend aus einem kognitiven Pretest und einem Standard-Pretest, mit Patientinnen und Patienten der Zielpopulation erprobt. Ziel des kognitiven Pretests ist die Untersuchung der Verständlichkeit und Interpretation der Fragen sowie des Zustandekommens der Antworten bei den Befragten. So können mögliche Probleme bei der Beantwortung der Fragen ermittelt und diese durch eine Überarbeitung des Fragebogenentwurfs ausgeräumt werden. Zur Prüfung der messtheoretischen Eigenschaften des überarbeiteten Fragebogens wird im Anschluss ein Standard-Pretest durchgeführt, wobei der Fragebogen unter ähnlichen Bedingungen,

wie sie im Regelbetrieb des jeweiligen QS-Verfahrens erwartet werden, in einer Patientenstichprobe eingesetzt wird. Bei der Umstellung auf eine Mixed-Mode-Befragung sollten sowohl der kognitive Pretest als auch der Standard-Pretest außerdem genutzt werden, um zu prüfen, ob die unterschiedlichen Erhebungsmodi vergleichbare Befragungsergebnisse liefern und die kombinierte Auswertung zulässig bzw. valide ist. Auf Grundlage des validierten Fragebogens werden schließlich die Qualitätsindikatoren definiert sowie Konzepte zur Datenerfassung und -auswertung erstellt.

Die einzelnen Entwicklungsschritte von Patientenbefragungen sowie entsprechende Stellen mit Anpassungsbedarf im Falle der Etablierung von Mixed-Mode-Befragungen sind in Abbildung 19 dargestellt.

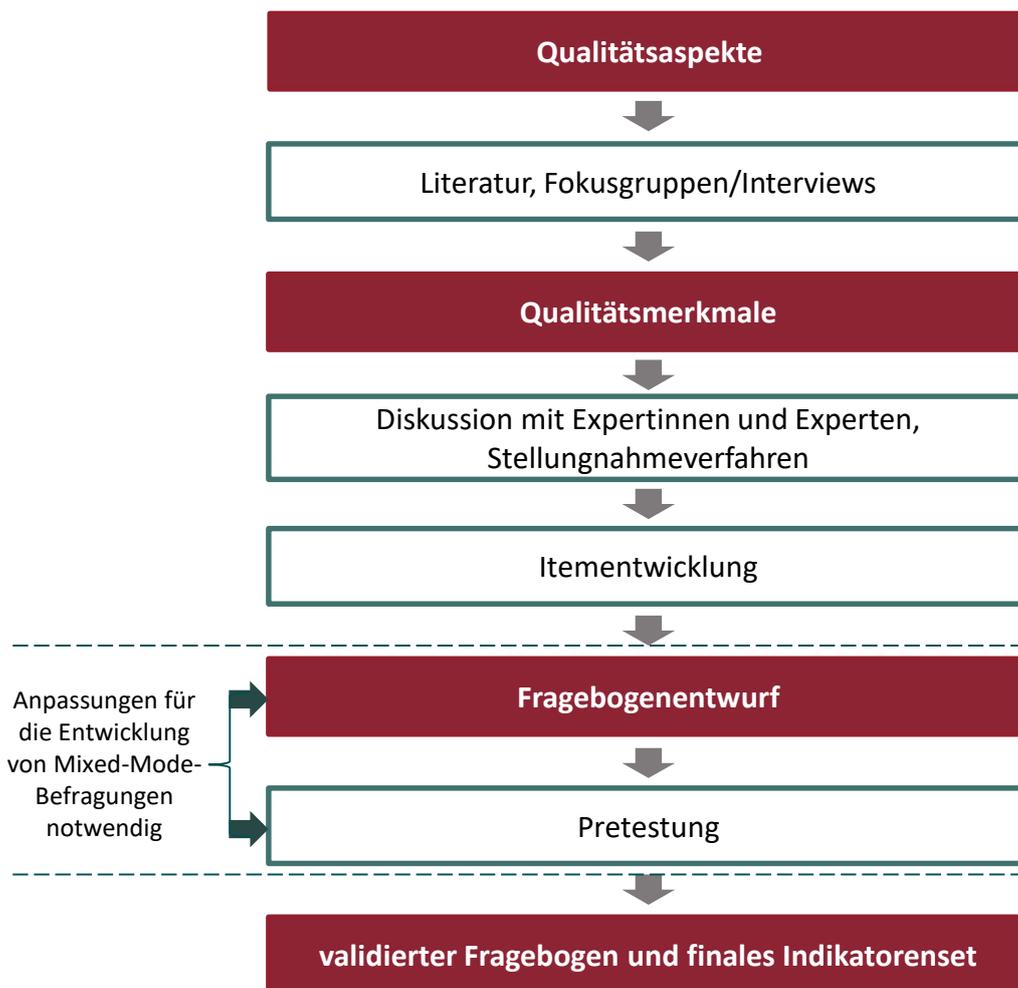


Abbildung 19: Entwicklungsschritte einer Patientenbefragung und entsprechende Stellen mit Anpassungsbedarf im Fall der Etablierung von Mixed-Mode-Befragungen

Im Folgenden sollen die Stellen im Entwicklungsprozess, für die sich im Falle von Mixed-Mode-Befragungen relevante Änderungen ergeben, kurz im Einzelnen dargestellt werden:

Erstellung des Fragebogens

- Zusätzlich zur Papierversion muss eine Onlineversion der Fragebögen erstellt werden, die insbesondere bereits im Rahmen des kognitiven Pretests und des Standard-Pretests eingesetzt

werden kann. Die Onlinebefragung muss so konzipiert werden, dass sie auf PC, Tabletcomputer und Smartphone unter Verwendung gängiger Webbrowser durchgeführt werden kann und dabei den geltenden Datenschutzstandards entspricht.

- Neben allgemeinen Empfehlungen zur Formulierung und Darstellung von Items sowie zum Layout von Fragebögen müssen spezifische Designmerkmale von Onlinefragebögen bei der Erstellung der Onlineversion berücksichtigt werden (vgl. Abschnitt 7.2).
- Gleichzeitig müssen die in Abschnitt 7.2.1 dargestellten Hinweise in Bezug auf das Design bei Mixed-Mode-Befragungen beachtet werden, um Modus-Effekte nach Möglichkeit zu vermeiden. Insbesondere soll sich die grafische Darstellung der Items nicht zu stark zwischen Papier- und Onlinefragebogen unterscheiden, sondern die Layouts der unterschiedlichen Befragungsmodi sollen nach Möglichkeit aufeinander abgestimmt werden.

Zielgruppenanalyse mit Hinblick auf Möglichkeit zur Push-to-Web Befragung

- Bereits vor der Phase der Pretestung sollte festgelegt werden, ob für die jeweilige Zielpopulation des QS-Verfahrens ein Push-to-Web Ansatz zielführend erscheint (vgl. Abschnitt 10.1.2). Hierzu bedarf es einer genauen Analyse der Zielpopulation mit Hinblick darauf, ob alle Subgruppen mit einem sequenziellen Mixed-Mode-Ansatz erreicht werden können und daher die Zielpopulation repräsentativ abgebildet werden kann.
- Dies ist vorrangig bei jungen, höher gebildeten und wenig multimorbiden bzw. im Alltag wenig eingeschränkten Patientengruppen der Fall (vgl. Kapitel 6).

Kognitiver Pretest

- Im kognitiven Pretest sollen die entwickelten Items sowohl im Papierformat als auch im Onlineformat getestet werden. Der Fokus sollte dabei über die bisherigen Ziele des kognitiven Pretests hinaus auch auf potenzielle Moduseffekte sowie auf spezifische Fehlerquellen aufgrund von zuvor getroffenen Entscheidungen zum Layout der onlinebasierten Befragung gelegt werden.
- Es ist zu beachten, dass es durch die Erweiterung der Zielsetzung des kognitiven Pretests zur Erhöhung der notwendigen Teilnehmeranzahl kommen kann.

Standard-Pretest

- Der Standard-Pretest kann bei entsprechend hohen Fallzahlen zur Überprüfung genutzt werden, ob die unterschiedlichen Erhebungsmodi zu vergleichbaren Daten führen und somit die kombinierte Auswertung valide Befragungsergebnisse erwarten lässt.

11.2 Überführung bereits bestehender papierbasierter Patientenbefragungen in eine Onlinebefragung

Entsprechend der Beauftragung durch den G-BA sollte auf Basis des vorliegenden Projektes erarbeitet werden, wie bereits bestehende oder in der Entwicklung befindliche Patientenbefragungen in eine Onlinebefragung überführt werden können. Dies betrifft zum einen die Identifikation von Anforderungen an Layout und Design von Onlinebefragungen, die auf bestehenden Papierfragebögen basieren (Zielsetzung II.a.). Zum anderen erfolgte eine Darstellung der notwendigen ergänzenden Entwicklungs- bzw. Validierungsschritte bei der Erstellung einer Onlineversion von bereits entwickelten Papierfragebögen (Zielsetzung II.b.).

Grundsätzlich treffen die in Abschnitt 11.1 genannten Punkte zu Neuentwicklungen von Mixed-Mode-Befragungen auch auf bereits vorliegende oder sich in Entwicklung befindliche Patientenbefragungen zu. Der Unterschied besteht darin, dass die Papierfragebögen im letztgenannten Fall in aller Regel bereits in ihrer grundlegenden Struktur vorliegen und diese daher retrospektiv nicht flexibel mit spezifischen Designüberlegungen zur Onlineversion abgestimmt werden können. Sobald ein Papierfragebogen gezielt für eine postalische Paper-Pencil-Befragung erstellt und entsprechend designt wurde und die Validierung im Rahmen von kognitivem Pretest und/oder Standard-Pretest eingeleitet bzw. bereits abschließend durchgeführt wurde, kann dieser Fragebogen nachträglich nur noch unter großem Aufwand verändert werden. Dies trifft sowohl auf bereits bestehende, aber auch auf sämtliche aktuell in Entwicklung befindlichen Patientenbefragungen des IQTIG zu. Für den Fall, dass die Überführung bereits entwickelter oder in Entwicklung befindlicher Patientenbefragungen in ein Onlineformat vom G-BA beauftragt wird, um eine Mixed-Mode-Befragung zu ermöglichen, muss sich somit die zu erstellende Onlineversion an den konkreten Ausgestaltungen des vorliegenden Papierfragebogens orientieren.

Wie in Abschnitt 11.1 dargestellt, ist für die Entwicklung einer Mixed-Mode-Befragung neben einer Zielgruppenanalyse mit Hinblick auf die Möglichkeit zur Push-to-Web-Befragung und der Konzeption und Erstellung der Onlinebefragung auch eine ausführliche Pretestung notwendig. Die Überführung bereits bestehender oder in der Entwicklung befindlicher Patientenbefragungen in eine Onlinebefragung erfordert daher in jedem Fall eine eigene Weiterentwicklung.

Literatur

- ADM [Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute] (2020): Jahresbericht 2019. [Stand:] 01.08.2020. Berlin: ADM. URL: https://www.adm-ev.de/wp-content/uploads/2020/09/ADM_Jahresbericht_2019_020920_WEB.pdf (abgerufen am: 05.08.2021).
- Alfonsson, S; Maathz, P; Hursti, T (2014): Interformat Reliability of Digital Psychiatric Self-Report Questionnaires: A Systematic Review. *Journal of Medical Internet Research* 16(12): e268. DOI: 10.2196/jmir.3395.
- Antoun, C; Couper, MP; Conrad, FG (2017): Effects of Mobile Versus PC Web on Survey Response Quality: A Crossover Experiment in a Probability Web Panel. *Public Opinion Quarterly* 81(S1): 280-306. DOI: 10.1093/poq/nfw088.
- Antoun, C; Katz, J; Argueta, J; Wang, L (2018): Design Heuristics for Effective Smartphone Questionnaires. *Social Science Computer Review* 36(5): 557-574. DOI: 10.1177/0894439317727072.
- Antoun, C; Conrad, FG; Couper, MP; West, BT (2019): Simultaneous Estimation of Multiple Sources of Error in a Smartphone-Based Survey. *Journal of Survey Statistics and Methodology* 7(1): 93-117. DOI: 10.1093/jssam/smy002.
- Bandilla, W (2015): Online – Befragungen. Version 1.1. [Stand:] Januar 2015. (GESIS Survey Guidelines). Mannheim: GESIS [Leibniz-Institut für Sozialwissenschaften]. DOI: 10.15465/gesis-sg_003.
- Bates, N; García Trejo, YA; Vines, M (2019): Are Sexual Minorities Hard-to-Survey? Insights from the 2020 Census Barriers, Attitudes, and Motivators Study (CBAMS) Survey. *Journal of Official Statistics* 35(4): 709-729. DOI: 10.2478/jos-2019-0030.
- Biemer, PP; de Leeuw, ED; Eckman, S; Edwards, B; Kreuter, F; Lyberg, LE; et al.; Hrsg. (2017): Total Survey Error in Practice Hoboken, US-NJ: Wiley. ISBN: 978-1-119-04167-2.
- Biffignandi, S; Bethlehem, J (2021): Handbook of Web Surveys. Second Edition. Hoboken, US-NJ: Wiley. ISBN: 978-1-119-37168-7.
- Blom, AG; Gathmann, C; Krieger, U (2015): Setting Up an Online Panel Representative of the General Population: The German Internet Panel. *Field Methods* 27(4): 391-408. DOI: 10.1177/1525822X15574494.
- Blumenberg, C; Barros, AJD (2018): Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature. *International Journal of Public Health* 63(6): 765-773. DOI: 10.1007/s00038-018-1108-4.
- BMBF [Bundesministerium für Bildung und Forschung] (2013): Glossar. Begriffe A-Z. ISCED 97. Bonn: BMBF. URL: <https://www.datenportal.bmbf.de/portal/de/G54.html> (abgerufen am: 15.07.2021).

- Börkan, B (2010): The Mode Effect in Mixed-Mode Surveys. Mail and Web Surveys. *Social Science Computer Review* 28(3): 371-380. DOI: 10.1177/0894439309350698.
- Börsch-Supan, A; Brandt, M; Hunkler, C; Kneip, T; Korbmacher, J; Malter, F; et al. (2013): Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology* 42(4): 992-1001. DOI: 10.1093/ije/dyt088.
- Börsch-Supan, A (2019): Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 7.1.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.710.
- Bosnjak, M; Dannwolf, T; Enderle, T; Schaurer, I; Struminskaya, B; Tanner, A; et al. (2017): Establishing an Open Probability-Based Mixed-Mode Panel of the General Population in Germany: The GESIS Panel. *Social Science Computer Review* 36(1): 103-115. DOI: 10.1177/0894439317697949.
- Bucks, B; Couper, MP; Fulford, SL (2020): A Mixed-Mode and Incentive Experiment Using Administrative Data. *Journal of Survey Statistics and Methodology* 8(2): 352-369. DOI: 10.1093/jssam/smz005.
- Callegaro, M; Lozar Manfreda, K; Vehovar, V (2015): Web Survey Methodology. Los Angeles, US-CA: Sage. ISBN: 978-0-85702-861-7.
- Campbell, N; Ali, F; Finlay, AY; Salek, SS (2015): Equivalence of electronic and paper-based patient-reported outcome measures. *Quality of Life Research* 24(8): 1949-1961. DOI: 10.1007/s11136-015-0937-3.
- Cohen, J (1988): Statistical Power Analysis for the Behavioral Sciences. Second Edition. New York, US-NY [u. a.]: Psychology Press. ISBN: 978-0-8058-0283-2.
- Conrad, FG; Couper, MP; Tourangeau, R; Zhang, C (2017): Reducing speeding in web surveys by providing immediate feedback. *Survey Research Methods* 11(1): 45-61. DOI: 10.18148/srm/2017.v11i1.6304.
- Cornesse, C; Bosnjak, M (2018): Is there an association between survey characteristics and representativeness? A meta-analysis. *Survey Research Methods* 12(1): 1-13. DOI: 10.18148/srm/2018.v12i1.7205.
- Couper, MP (2008): Designing Effective Web Surveys. Cambridge, US-MA: Cambridge University Press. ISBN: 978-0-521-71794-6.
- Couper, MP; Miller, PV (2008): Web Survey Methods. Introduction. *Public Opinion Quarterly* 72(5): 831-835. DOI: 10.1093/poq/nfn066.
- Couper, MP (2011): The Future of Modes of Data Collection. *Public Opinion Quarterly* 75(5): 889-908. DOI: 10.1093/poq/nfr046.
- Couper, MP; Antoun, C; Mavletova, A (2017): Mobile Web Surveys. Section 2, Chapter 7. In: Biemer, PP; Leeuw, E; Eckman, S; Edwards, B; Kreuter, F; Lyberg, LE; et al.: *In Total Survey Error in Practice*. Hoboken, US-NJ: Wiley, 133-154. ISBN: 978-1-119-04167-2.

- Daikeler, J; Bošnjak, M; Lozar Manfreda, K (2020): Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. *Journal of Survey Statistics and Methodology* 8(3): 513-539. DOI: 10.1093/jssam/smz008.
- De Bruijne, M; Wijnant, A (2014): Improving Response Rates and Questionnaire Design for Mobile Web Surveys. *Public Opinion Quarterly* 78(4): 951-962. DOI: 10.1093/poq/nfu046.
- de Leeuw, ED; Hox, JJ (2010): Internet Surveys as Part of a Mixed-Mode Design. Chapter 3. In: Das, M; Ester, P; Kaczmirek, L; Hrsg.: *Social and Behavioral Research and the Internet. Advances in Applied Methods and Research Strategies*. New York, US-NY: Routledge, 45-76. ISBN: 978-1-84872-817-2.
- Debell, M; Wilson, C; Jackman, S; Figueroa, L (2021): Optimal Response Formats for Online Surveys: Branch, Grid, or Single Item? *Journal of Survey Statistics and Methodology* 9(1): 1-24. DOI: 10.1093/jssam/smz039.
- DeLeeuw, ED (2018): Mixed-Mode: Past, Present, and Future. *Survey Research Methods* 12(2): 75-89. DOI: 10.18148/srm/2018.v12i2.7402.
- Destatis [Statistisches Bundesamt] (2021a): Erhebung über die private Nutzung von Informations- und Kommunikationstechnologien. IKT 2020. [Stand:] 04.02.2021. (Qualitätsbericht). Wiesbaden: Destatis. URL: <https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Einkommen-Konsum-Lebensbedingungen/ikt-private-haushalte-2020.pdf> [Download: Herunterladen (PDF, 679KB, Datei ist nicht barrierefrei)] (abgerufen am: 05.08.2021).
- Destatis [Statistisches Bundesamt] (2021b): Wirtschaftsrechnungen. Private Haushalte in der Informationsgesellschaft – Nutzung von Informations- und Kommunikationstechnologien. 2020. [Stand:] 20.01.2021. (Fachserie 15, Reihe 4). Wiesbaden: Destatis. Artikelnummer: 2150400207004. URL: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/Publikationen/Downloads-IT-Nutzung/private-haushalte-ikt-2150400207004.pdf> [Download: Herunterladen (PDF, 850KB, Datei ist nicht barrierefrei)] (abgerufen am: 05.08.2021).
- Dillman, D (2017): The promise and challenge of pushing respondents to the web in mixed-mode surveys. *Survey Methodology* 43(1): 3-30. Statistics Canada, Catalogue No. 12-001-X. URL: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2017001/article/14836-eng.pdf> (abgerufen am: 20.01.2021).
- Dillman, DA; Smyth, JD; Christian, LM (2014a): Internet, Phone, Mail and Mixed-Mode Surveys. The Tailored Design Method. Fourth Edition. Hoboken, US-NJ: Wiley. ISBN: 978-1-118-45614-9.
- Dillman, DA; Smyth, JD; Christian, LM (2014b): Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, US-NJ: John Wiley & Sons. ISBN: 978-1-118-45614-9.

- Dodou, D; de Winter, JCF (2014): Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior* 36: 487-495. DOI: 10.1016/j.chb.2014.04.005.
- Dykema, J; Stevenson, J; Klein, L; Kim, Y; Day, B (2012): Effects of E-Mailed Versus Mailed Invitations and Incentives on Response Rates, Data Quality, and Costs in a Web Survey of University Faculty. *Social Science Computer Review* 31(3): 359-370. DOI: 10.1177/0894439312465254.
- Efkeky, A; Gillies, K; Gardner, H; Fraser, C; Ishaku, T; Treweek, S (2020): Non-randomised evaluations of strategies to increase participant retention in randomised controlled trials: a systematic review. *Systematic Reviews* 9:224. DOI: 10.1186/s13643-020-01471-x.
- ESS [European Social Survey] Round 9 (2021): European Social Survey Round 9 Data. Data file edition 3.1. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. DOI: 10.21338/NSD-ESS9-2018.
- Eurostat (2021): Bevölkerung am 1. Januar nach Altersgruppe und Geschlecht. [Stand:] 05.07.2021. Brüssel, BE: Europäische Kommission. URL: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_pjangroup&lang=de (abgerufen am: 15.07.2021).
- Flanagan, SM; Greenfield, S; Coad, J; Neilson, S (2015): An exploration of the data collection methods utilised with children, teenagers and young people (CTYPs). *BMC: Research Notes* 8:61. DOI: 10.1186/s13104-015-1018-y.
- Freedman, VA; McGonagle, KA; Couper, MP (2018): Use of a Targeted Sequential Mixed Mode Protocol in a Nationally Representative Panel Study. *Journal of Survey Statistics and Methodology* 6(1): 98-121. DOI: 10.1093/jssam/smx012.
- G-BA [Gemeinsamer Bundesausschuss] (2020): Beschluss des Gemeinsamen Bundesausschusses über eine Beauftragung des IQTIG mit der Entwicklung eines Konzeptes zur onlinebasierten Patientenbefragung. [Stand:] 20.11.2020. Berlin: G-BA. URL: https://www.g-ba.de/downloads/39-261-4553/2020-11-20_IQTIG-Beauftragung-Patientenbefragung.pdf (abgerufen am: 07.12.2020).
- GESIS [Leibniz Institute für Sozialwissenschaften] (2013): ALLBUS/GGSS 2012 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2012). ZA4614 Datenfile Version 1.1.1. Köln: GESIS Datenarchiv. DOI: 10.4232/1.11753.
- GESIS [Leibniz Institute für Sozialwissenschaften] (2017): ZA5250: ALLBUS/GGSS 2016 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2016). ZA5250 Datenfile Version 2.1.0. Köln: GESIS Datenarchiv. DOI: 10.4232/1.12796.
- GESIS [Leibniz Institute für Sozialwissenschaften] (2018): ALLBUS/GGSS 2014 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2014). ZA5240 Datenfile Version 2.2.0. Köln: GESIS Datenarchiv. DOI: 10.4232/1.13141.

- GESIS [Leibniz Institute für Sozialwissenschaften] (2019): ZA5270: ALLBUS/GGSS 2018 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2018). ZA5270 Datenfile Version 2.0.0. Köln: GESIS Datenarchiv. DOI: 10.4232/1.13250.
- Gnamb, T; Kaspar, K (2017): Socially Desirable Responding in Web-Based Questionnaires: A Meta-Analytic Review of the Candor Hypothesis. *Assessment* 24(6): 746-762. DOI: 10.1177/1073191115624547.
- Greenleaf, AR; Gibson, DG; Khattar, C; Labrique, AB; Pariyo, GW (2017): Building the Evidence Base for Remote Data Collection in Low- and Middle-Income Countries: Comparing Reliability and Accuracy Across Survey Modalities. *Journal of Medical Internet Research* 19(5): e140. DOI: 10.2196/jmir.7331.
- Greszki, R; Meyer, M; Schoen, H (2015): Exploring the Effects of Removing “Too Fast” Responses and Respondents from Web Surveys. *Public Opinion Quarterly* 79(2): 471-503. DOI: 10.1093/poq/nfu058.
- Groves, RM; Fowler, Jr. F. J.; Couper, MP; Lepkowski, JM; Singer, E; Tourangeau, R (2009): *Survey Methodology* Second Edition. (Wiley Series in Survey Methodology). Hoboken, US-NJ: Wiley. ISBN: 978-0-470-46546-2.
- Gummer, T; Daikeler, J (2020): A Note on How Prior Survey Experience With Self-Administered Panel Surveys Affects Attrition in Different Modes. *Social Science Computer Review* 38(4): 490-498. DOI: 10.1177/0894439318816986.
- Haer, R; Meidert, N (2013): Does the first impression count? Examining the effect of the welcome screen design on the response rate. *Survey Methodology* 39(2): 419-434. Catalogue No. 12-001-X. URL: <https://www150.statcan.gc.ca/pub/12-001-x/12-001-x2013002-eng.pdf> (abgerufen am: 05.08.2021).
- Hu, J (2020): Horizontal or Vertical? The Effects of Visual Orientation of Categorical Response Options on Survey Responses in Web Surveys. *Social Science Computer Review* 38(6): 779-792. DOI: 10.1177/0894439319834296.
- Hussain, N; Sprague, S; Madden, K; Hussain, FN; Pindiprolu, B; Bhandari, M (2015): A Comparison of the Types of Screening Tool Administration Methods Used for the Detection of Intimate Partner Violence: A Systematic Review and Meta-Analysis. *Trauma, Violence & Abuse* 16(1): 60-69. DOI: 10.1177/1524838013515759.
- IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2018a): Entwicklung einer Befragung von Patienten und Patientinnen mit Schizophrenie, schizotypen und wahnhaften Störungen. Entwicklung einer Patientenbefragung im Rahmen der Aktualisierung und Erweiterung des QS-Verfahrens *Versorgung von volljährigen Patienten und Patientinnen mit Schizophrenie, schizotypen und wahnhaften Störungen*. Abschlussbericht. Stand: 15.12.2018. Berlin: IQTIG. URL: https://iqtig.org/downloads/berichte/2018/IQTIG_Patientenbefragung_QS-Verfahren-Schizophrenie_Abschlussbericht_2018-12-15_barrierefrei.pdf (abgerufen am: 28.10.2021).

- IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2018b):
Entwicklung von Patientenbefragungen im Rahmen des Qualitätssicherungsverfahrens
Perkutane Koronarintervention und Koronarangiographie. Abschlussbericht. Stand:
15.12.2018. Berlin: IQTIG. URL:
https://iqtig.org/downloads/berichte/2018/IQTIG_Patientenbefragung_QS-PCI_Abschlussbericht_2018-12-15_barrierefrei.pdf (abgerufen am: 19.03.2021).
- IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2019):
Methodische Grundlagen V1.1. Stand: 15.04.2019. Berlin: IQTIG. URL:
https://iqtig.org/dateien/dasiqtig/grundlagen/IQTIG_Methodische-Grundlagen-V1.1_barrierefrei_2019-04-15.pdf (abgerufen am: 21.10.2021).
- IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2020):
Entwicklung von Patientenbefragungen für das Qualitätssicherungsverfahren
Nierenersatztherapie. Abschlussbericht. Stand: 02.11.2020. Berlin: IQTIG.
[unveröffentlicht].
- Jibb, LA; Khan, JS; Seth, P; Laloo, C; Mulrooney, L; Nicholson, K; et al. (2020): Electronic Data Capture Versus Conventional Data Collection Methods in Clinical Pain Studies: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research* 22(6): e16480. DOI: 10.2196/16480.
- Kernan, WN; Viscoli, CM; Makuch, RW; Brass, LM; Horwitz, RI (1999): Stratified Randomization for Clinical Trials. *Journal of Clinical Epidemiology* 52(1): 19-26. DOI: 10.1016/S0895-4356(98)00138-3.
- Keusch, F; Yan, T (2017): Web Versus Mobile Web: An Experimental Study of Device Effects and Self-Selection Effects. *Social Science Computer Review* 35(6): 751-769. DOI: 10.1177/0894439316675566.
- Lenzner, T; Kaczmirek, L; Galesic, M (2014): Left Feels Right: A Usability Study on the Position of Answer Boxes in Web Surveys. *Social Science Computer Review* 32(6): 743-764. DOI: 10.1177/0894439313517532.
- Liu, M; Cernat, A (2016): Item-by-item Versus Matrix Questions: A Web Survey Experiment. *Social Science Computer Review* 36(6): 690-706. DOI: 10.1177/0894439316674459.
- Liu, M; Wronski, L (2018): Examining Completion Rates in Web Surveys via Over 25,000 Real-World Surveys. *Social Science Computer Review* 36(1): 116-124. DOI: 10.1177/0894439317695581.
- Lugtig, P; Toepoel, V (2016): The Use of PCs, Smartphones, and Tablets in a Probability-Based Panel Survey: Effects on Survey Measurement Error. *Social Science Computer Review* 34(1): 78-94. DOI: 10.1177/0894439315574248.
- Malter, F; Börsch-Supan, A; Hrsg. (2017): SHARE Wave 6: Panel innovations and collecting Dried Blood Spots. München: MEA [Munich Center for the Economics of Aging], Max-Planck-Instituts für Sozialrecht und Sozialpolitik. ISBN: 978-3-00-058556-2. URL: <http://www.share->

project.org/fileadmin/pdf_documentation/MFRB_SHARE_Wave_6_Panel_innovation/2018_04_SHARE-WAVE-6_MFRB.pdf (abgerufen am: 05.08.2021).

- Marcano Belisario, JS; Jamsek, J; Huckvale, K; O'Donoghue, J; Morrison, CP; Car, J (2015): Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods (Review). *Cochrane Database of Systematic Reviews* (7): Art. No.: MR000042. DOI: 10.1002/14651858.MR000042.pub2.
- Mavletova, A; Couper, MP (2016): Grouping of Items in Mobile Web Questionnaires. *Field Methods* 28(2): 170-193. DOI: 10.1177/1525822X15595151.
- Mavletova, AM; Couper, MP (2015): A Meta-Analysis of Breakoff Rates in Mobile Web Surveys. Chapter 6. In: Toninelli, D; Pinter, R; de Pedraza, P: *Mobile Research Methods. Opportunities and challenges of mobile research methodologies*. London, GB: Ubiquity Press, 81-98. DOI: 10.5334/bar.
- Mavletova, AM; Couper, MP; Lebedev, D (2017): Grid and Item-by-Item Formats in PC and Mobile Web Surveys. *Social Science Computer Review* 36(6): 647-668. DOI: 10.1177/0894439317735307.
- McMaster, HS; LeardMann, CA; Speigle, S; Dillman, DA (2017): An experimental comparison of web-push vs. paper-only survey procedures for conducting an in-depth health survey of military spouses. *BMC: Medical Research Methodology* 17:73. DOI: 10.1186/s12874-017-0337-1.
- 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. DOI: 10.1093/poq/nfs047.
- Meirte, J; Hellemans, N; Anthonissen, M; Denteneer, L; Maertens, K; Moortgat, P; et al. (2020): Benefits and Disadvantages of Electronic Patient-reported Outcome Measures: Systematic Review. *JMIR Perioperative Medicine* 3(1): e15588. DOI: 10.2196/15588.
- Messer, BL; Dillman, DA (2011): Surveying the General Public over the Internet Using Address-Based Sampling and Mail Contact Procedures. *Public Opinion Quarterly* 75(3): 429-457. DOI: 10.1093/poq/nfr021.
- Millar, MM; Dillman, DA (2011): Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly* 75(2): 249-269. DOI: 10.1093/poq/nfr003.
- Muehlhausen, W; Doll, H; Quadri, N; Fordham, B; O'Donohoe, P; Dogar, N; et al. (2015): Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health and Quality of Life Outcomes* 13:167. DOI: 10.1186/s12955-015-0362-x.
- Page, MJ; Moher, D; Bossuyt, PM; Boutron, I; Hoffmann, TC; Mulrow, CD; et al. (2021): PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ: Research Methods & Reporting* 372:n160. DOI: 10.1136/bmj.n160.

- Patrick, ME; Couper, MP; Parks, MJ; Laetz, V; Schulenberg, JE (2020): Comparison of a web-push survey research protocol with a mailed paper and pencil protocol in the Monitoring the Future panel survey. *Addiction* 116(1): 191-199. DOI: 10.1111/add.15158.
- Pieper, D; Antoine, S-L; Mathes, T; Neugebauer, EAM; Eikermann, M (2014): Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology* 67(4): 368-375. DOI: 10.1016/j.jclinepi.2013.11.007.
- Pollock, M; Fernandes, RM; Becker, LA; Pieper, D; Hartling, L (2020): Overviews of Reviews. Part 1, Chapter V. In: Higgins, JPT; Thomas, J; Chandler, J; Cumpston, M; Li, T; Page, MJ; et al.; Hrsg.: *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.1. Updated September 2020. Cochrane. URL: <https://training.cochrane.org/handbook/current/chapter-v> (abgerufen am: 30.11.2020).
- R Core Team (2021): R: A Language and Environment for Statistical Computing [*Open Source Software*]. R version 4.1.1 (Kick Things). Wien, AT: R Foundation for Statistical Computing. URL: <https://cran.r-project.org/src/base/R-4/> [Download: R-4.0.5.tar.gz] (abgerufen am: 14.09.2021).
- Research], ATAAfPO (2016): Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. [Washington, US-DC]: AAPOR. URL: https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf (abgerufen am: 13.09.2021).
- Revilla, M; Couper, MP (2018): Comparing Grids With Vertical and Horizontal Item-by-Item Formats for PCs and Smartphones. *Social Science Computer Review* 36(3): 349-368. DOI: 10.1177/0894439317715626.
- Rutherford, C; Costa, D; Mercieca-Bebber, R; Rice, H; Gabb, L; King, M (2016): Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. *Quality of Life Research* 25(3): 559-574. DOI: 10.1007/s11136-015-1110-8.
- Sakshaug, JW; Vicari, B; Couper, MP (2018): Paper, E-mail, or Both? Effects of Contact Mode on Participation in a Web Survey of Establishments. *Social Science Computer Review* 37(6): 750-765. DOI: 10.1177/0894439318805160.
- Schlosser, S; Mays, A (2017): Mobile and Dirty: Does Using Mobile Devices Affect the Data Quality and the Response Process of Online Surveys? *Social Science Computer Review* 36(2): 212-230. DOI: 10.1177/0894439317698437.
- Schnell, R; Hill, PB; Esser, E (2018): Methoden der empirischen Sozialforschung. 11., überarbeitete Auflage. Berlin [u. a.]: De Gruyter Oldenbourg. ISBN: 978-3-11-057732-7.
- Schnell, R (2019): Survey-Interviews. Methoden standardisierter Befragungen. 2. Auflage. (Studienskripten zur Soziologie). Wiesbaden: Springer. ISBN: 978-3-531-19900-9.
- Schonlau, M; Couper, MP (2017): Options for Conducting Web Surveys. *Statistical Science* 32(2): 279-292. DOI: 10.1214/16-STS597.

- SHARE [Survey of Health, Ageing and Retirement in Europe-] – ERIC [European Research Infrastructure Consortium], (2021): Citation Requirements. SHARE acknowledgement(s). München: SHARE-ERIC. URL: <http://www.share-project.org/data-access/citation-requirements.html> (abgerufen am: 05.08.2021).
- Shea, BJ; Reeves, BC; Wells, G; Thuku, M; Hamel, C; Moran, J; et al. (2017): AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ: Research Methods & Reporting* 358:j4008. DOI: 10.1136/bmj.j4008.
- Shin, E; Johnson, TP; Rao, K (2012): Survey Mode Effects on Data Quality: Comparison of Web and Mail Modes in a U.S. National Panel Survey. *Social Science Computer Review* 30(2): 212-228. DOI: 10.1177/0894439311404508.
- SIGN [Scottish Intercollegiate Guidelines Network] ([kein Datum]): [Search filters:] Systematic Reviews. Edinburgh, GB: SIGN. URL: <https://www.sign.ac.uk/assets/search-filters-systematic-reviews.docx> (abgerufen am: 21.09.2020).
- Sommer, J; Diedenhofen, B; Musch, J (2016): Not to Be Considered Harmful: Mobile-Device Users Do Not Spoil Data Quality in Web Surveys. *Social Science Computer Review* 35(3): 378-387. DOI: 10.1177/0894439316633452.
- Stadtmüller, S; Beuthner, C; Silber, H (2021): Mixed-Mode Surveys. Version 1.0. [Stand:] January 2021. (GESIS- Survey Guidelines). Mannheim: GESIS [Leibniz Institute für Sozialwissenschaften]. DOI: 10.15465/gesis-sg_en_038.
- Stadtmüller, S; Silber, H (2021): Selbstadministrierte Mixed-Mode-Befragungen: Die Zukunft der Datenerhebung? Meet the Experts. Best-Practice-Methoden in der Umfrageforschung. [Stand:] 28.01.2021. Köln: GESIS [Leibniz Institute für Sozialwissenschaften]. URL: https://www.gesis.org/fileadmin/user_upload/MeettheExperts/GESIS_MeettheExperts_mixed-mode.pdf (abgerufen am: 05.08.2021).
- Toepoel, V; Lugtig, P (2014): What Happens if You Offer a Mobile Option to Your Web Panel? Evidence From a Probability-Based Panel of Internet Users. *Social Science Computer Review* 32(4): 544-560. DOI: 10.1177/0894439313510482.
- Toninelli, D; Revilla, M (2016): Smartphones vs PCs: Does the Device Affect the Web Survey Experience and the Measurement Error for Sensitive Topics? A Replication of the Mavletova & Couper's 2013 Experiment. *Survey Research Methods* 10(2): 153-169. DOI: 10.18148/srm/2016.v10i2.6274.
- Tourangeau, R; Couper, MP; Conrad, FG (2013): "Up Means Good". The Effect of Screen Position on Evaluative Ratings in Web Surveys. *Public Opinion Quarterly* 77(S1): 69-88. DOI: 10.1093/poq/nfs063.
- Tourangeau, R; Maitland, A; Rivero, G; Sun, H; Williams, D; Yan, T (2017a): Web Surveys by Smartphone and Tablets. Effects on Survey Responses. *Public Opinion Quarterly* 81(4): 896-929. DOI: 10.1093/poq/nfx035.

- Tourangeau, R; Sun, H; Yan, T; Maitland, A; Rivero, G; Williams, D (2017b): Web Surveys by Smartphones and Tablets: Effects on Data Quality. *Social Science Computer Review* 36(5): 542-556. DOI: 10.1177/0894439317719438.
- Vannieuwenhuyze, J (2014): On the Relative Advantage of Mixed-Mode versus Single-Mode Surveys. *Survey Research Methods* 8(1): 31-42. DOI: 10.18148/srm/2014.v8i1.5500.
- Vicente, P; Reis, E (2010): Using Questionnaire Design to Fight Nonresponse Bias in Web Surveys. *Social Science Computer Review* 28(2): 251-267. DOI: 10.1177/0894439309340751.
- Villar, A; Callegaro, M; Yang, Y (2013): Where Am I? A Meta-Analysis of Experiments on the Effects of Progress Indicators for Web Surveys. *Social Science Computer Review* 31(6): 744-762. DOI: 10.1177/0894439313497468.
- Volland, C; Rosumeck, S; Wagner, M; Blatt, K; Euler, U; Liegl, G (2020): Does mode of data collection affect the results of self-administered surveys? An overview of systematic reviews [*Study protocol*]. [Stand:] 29.01.2021. Berlin: IQTIG [Institut für Qualität und Transparenz im Gesundheitswesen]. URL: <https://osf.io/f748j/download> (abgerufen am: 01.02.2021).
- VuMa [Arbeitsgemeinschaft Verbrauchs- und Medienanalyse] Touchpoints (2021): Konsumenten im Fokus. Basisinformationen für fundierte Mediaentscheidungen. VuMA Touchpoints 2021. Frankfurt am Main [u. a.]: VuMa. URL: https://www.vuma.de/fileadmin/user_upload/PDF/berichtsbaende/VuMA_Berichtsband_2_021.pdf (abgerufen am: 05.08.2021).
- Waffenschmidt, S; Janzen, T; Hausner, E; Kaiser, T (2013): Simple search techniques in PubMed are potentially suitable for evaluating the completeness of systematic reviews. *Journal of Clinical Epidemiology* 66(6): 660-665. DOI: 10.1016/j.jclinepi.2012.11.011.
- Weigold, A; Weigold, K; Natera, SN (2018a): Mean scores for self-report surveys completed using paper-and-pencil and computers: A meta-analytic test of equivalence. *Computers in Human Behavior* 86: 153-164. DOI: 10.1016/j.chb.2018.04.038.
- Weigold, A; Weigold, IK; Natera, SN (2018b): Response Rates for Surveys Completed With Paper-and-Pencil and Computers: Using Meta-Analysis to Assess Equivalence. *Social Science Computer Review* 37(5): 649-668. DOI: 10.1177/0894439318783435.
- Wells, T; Bailey, JT; Link, MW (2013): Comparison of Smartphone and Online Computer Survey Administration. *Social Science Computer Review* 32(2): 238-255. DOI: 10.1177/0894439313505829.
- Wenz, A; Jäckle, A; Couper, MP (2019): Willingness to use mobile technologies for data collection in a probability household panel. *Survey Research Methods* 13(1): 1-22. DOI: 10.18148/srm/2019.v1i1.7298.
- White, MK; Maher, SM; Rizio, AA; Bjorner, JB (2018): A meta-analytic review of measurement equivalence study findings of the SF-36® and SF-12® Health Surveys across electronic

modes compared to paper administration. *Quality of Life Research* 27(7): 1757-1767. DOI: 10.1007/s11136-018-1851-2.

Yan, T; Conrad, FG; Tourangeau, R; Couper, MP (2011): Should I Stay or Should I go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys. *International Journal of Public Opinion Research* 23(2): 131-147. DOI: 10.1093/ijpor/edq046.

Anhang

Anhang 1: Suchstrategien der systematischen Literaturrecherche.....	133
Anhang 2: Anwendung des AMSTAR 2-Instruments im vorliegenden Projekt	140
Anhang 3: Ausgeschlossene Artikel in der systematischen Literaturrecherche	142
Anhang 4: Charakteristika der eingeschlossenen systematischen Reviews	146
Anhang 5: AMSTAR-2-Bewertung der eingeschlossenen systematischen Reviews	150
Anhang 6: Ergebnisse der systematischen Literaturrecherche: Response-Rate	152
Anhang 7: Ergebnisse der systematischen Literaturrecherche: Antwortverhalten – Vergleiche mit Paper-Pencil-Befragung	154
Anhang 8: Ergebnisse der systematischen Literaturrecherche: Antwortverhalten – Vergleiche elektronischer Erhebungsmodi	163
Anhang 9: Unit-Nonresponse-Analyse im Rahmen der empirischen Untersuchung.....	166
Anhang 10: Analyse der „Weiß nicht mehr“-Antworten im Rahmen der empirischen Untersuchung.....	167

Anhang 1: Suchstrategien der systematischen Literaturrecherche

Tabelle A 1: Suchstrategie für MEDLINE via Ovid; Datum der Recherche: 01.02.2021

#	Searches
1	Meta-Analysis as Topic/ or exp Review Literature as Topic/
2	(meta analy* or metaanaly*).tw.
3	Meta-Analysis/ or Systematic Review/
4	((systematic* or comprehensive* or integrative or electronic*) adj3 (review? or overview? or bibliographic* or literature or search*)).tw.
5	1 or 2 or 3 or 4
6	(medline or pubmed or cochrane or embase or psychinfo or psycinfo or cinahl or cinhal or science citation index or scopus or science?direct or google scholar).ab.
7	((study or studies) adj (selection or selected)).ab.
8	(reference list* or bibliograph* or hand-search* or handsearch* or relevant journals or manual search*).ab.
9	6 or 7 or 8
10	(selection criteria or data extraction).ab.
11	Review/
12	10 and 11
13	5 or 9 or 12
14	limit 13 to yr="2011 -Current"
15	limit 14 to (english or german)
16	exp Cell Phone/
17	exp Internet/
18	exp Correspondence as Topic/
19	Reminder Systems/
20	Minicomputers/ or Computers, Handheld/
21	Mobile Applications/
22	Paper/
23	((mode or modes or method? or way?) adj3 (administ* or deliver* or survey? or differen* or collect* or compar* or option? or choice? or conduct* or carr* or complet*)).tw.
24	(internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app? or application? or mobile or push or ipad or tablet or handwritten).tw.

#	Searches
25	16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24
26	exp *"Surveys and Questionnaires"/
27	*Data Collection/mt [Methods]
28	*Outcome Assessment, Health Care/mt [Methods]
29	exp Patient Reported Outcome Measures/
30	((questionnaire? or survey?) adj3 (internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app? or application? or mobile or push or ipad or tablet or handwritten)).tw.
31	((patient* or self) adj2 report* adj2 (outcome* or experien* or measure*)).tw.
32	(PROM? or PREM? or PRM? or PRO or ePRO?).tw.
33	26 or 27 or 28 or 29 or 30 or 31 or 32
34	15 and 25 and 33
35	remove duplicates from 34

Tabelle A 2: Suchstrategie für APA PsycInfo via Ovid; Datum der Recherche: 01.02.2021

#	Searches
1	(meta analy* or metaanaly*).tw.
2	exp Meta-Analysis/ or exp "Literature Review"/
3	((systematic* or comprehensive* or integrative or electronic*) adj3 (review? or overview? or bibliographic* or literature or search*)).tw.
4	1 or 2 or 3
5	(medline or pubmed or cochrane or embase or cinahl or cinhal or science citation index or scopus or science?direct or google scholar).ab.
6	((study or studies) adj (selection or selected)).ab.
7	(reference list* or bibliograph* or hand-search* or handsearch* or relevant journals or manual search*).ab.
8	4 or 5 or 6 or 7
9	limit 8 to yr="2011 -Current"
10	limit 9 to (english or german)
11	exp mobile phones/
12	exp internet/
13	exp nonverbal communication/

#	Searches
14	exp electronic communication/
15	exp Written Communication/
16	exp mobile devices/
17	exp Mobile Applications/
18	((mode or modes or method? or way?) adj3 (administ* or deliver* or survey? or differen* or collect* or compar* or option? or choice? or conduct* or carr* or complet*)).tw.
19	(internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app? or application? or mobile or push or ipad or tablet or handwritten).tw.
20	11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19
21	exp *questionnaires/
22	exp *surveys/
23	*Data Collection/
24	*health outcomes/
25	Patient Reported Outcome Measures/
26	((questionnaire? or survey?) adj3 (internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app? or application? or mobile or push or ipad or tablet or handwritten)).tw.
27	((patient* or self) adj2 report* adj2 (outcome* or experien* or measure*)).tw.
28	(PROM? or PREM? or PRM? or PRO or ePRO?).tw.
29	21 or 22 or 23 or 24 or 25 or 26 or 27 or 28
30	10 and 20 and 29
31	remove duplicates from 30

Suchstrategie für Embase via Elsevier

Tabelle A 3: Suchstrategie für Embase via Elsevier; Datum der Recherche: 01.02.2021

No.	Query
#1	'systematic review (topic)'/de OR 'systematic review'/de OR 'meta analysis (topic)'/de OR 'meta analysis'/exp
#2	metaanaly*:ab,ti OR 'meta analy*':ab,ti
#3	((systematic* OR comprehensive* OR integrative OR electronic*) NEAR/3 (review\$ OR overview\$ OR bibliographic* OR literature OR search*)):ab,ti
#4	#1 OR #2 OR #3

No.	Query
#5	medline:ab OR pubmed:ab OR cochrane:ab OR embase:ab OR psychinfo:ab OR psycinfo:ab OR cinahl:ab OR cinhal:ab OR 'science citation index':ab OR scopus:ab OR sciencedirect:ab OR 'science direct':ab OR 'google scholar':ab
#6	((study OR studies) NEAR/1 (selection OR selected)):ab
#7	'reference list*':ab OR bibliograph*':ab OR 'hand-search*':ab OR 'handsearch*':ab OR 'relevant journals':ab OR 'manual search*':ab
#8	#5 OR #6 OR #7
#9	'selection criteria':ab OR 'data extraction':ab
#10	'review'/de
#11	#9 AND #10
#12	#4 OR #8 OR #11
#13	#12 AND [2011-2021]/py
#14	#13 AND ([english]/lim OR [german]/lim)
#15	'mobile phone'/exp OR 'internet'/exp OR 'writing'/de OR 'minicomputer'/de OR 'personal digital assistant'/de OR 'mobile application'/exp OR 'paper'/de OR 'reminder system'/de
#16	((mode OR modes OR method\$ OR way\$) NEAR/3 (administ* OR deliver* OR survey\$ OR differen* OR collect* OR compar* OR option\$ OR choice\$ OR conduct* OR carr* OR complet*)):ti,ab
#17	internet:ti,ab OR web:ti,ab OR online:ti,ab OR electronic:ti,ab OR paper:ti,ab OR pen:ti,ab OR postal:ti,ab OR pencil:ti,ab OR email:ti,ab OR mail*:ti,ab OR app\$:ti,ab OR application\$:ti,ab OR mobile:ti,ab OR push:ti,ab OR ipad:ti,ab OR tablet:ti,ab OR handwritten:ti,ab
#18	#15 OR #16 OR #17
#19	'questionnaire'/exp/mj OR 'health survey'/exp/mj
#20	'outcome assessment'/mj OR 'data collection method'/mj
#21	'patient-reported outcome'/de
#22	((questionnaire\$ OR survey\$) NEAR/3 (internet OR web OR online OR electronic OR paper OR pen OR postal OR pencil OR email OR mail* OR app\$ OR application\$ OR mobile OR push OR ipad OR tablet OR handwritten)):ti,ab
#23	((patient* OR self) NEAR/2 report* NEAR/2 (outcome* OR experien* OR measure*)):ti,ab
#24	prom\$:ti,ab OR prem\$:ti,ab OR prm\$:ti,ab OR pro:ti,ab OR epro\$:ti,ab
#25	#19 OR #20 OR #21 OR #22 OR #23 OR #24
#26	#14 AND #18 AND #25

Tabelle A 4: Suchstrategie für CINAHL via EBSCO; Datum der Recherche: 01.02.2021

#	Query
S1	TI ("meta analy*" OR metaanaly*) OR AB ("meta analy*" OR metaanaly*)
S2	(MH "Meta Analysis") OR (MH "Systematic Review")
S3	TI ((systematic* or comprehensive* or integrative or electronic*) N3 (review# or overview# or bibliographic* or literature or search*)) OR AB ((systematic* or comprehensive* or integrative or electronic*) N3 (review# or overview# or bibliographic* or literature or search*))
S4	S1 OR S2 OR S3
S5	AB (medline OR pubmed OR cochrane OR embase OR psychinfo OR psycinfo OR cinahl OR cinhal OR sciencedirect OR scopus) OR AB ("science citation index") OR AB ("science direct") OR AB ("google scholar")
S6	AB ((study or studies) N1 (selection or selected))
S7	AB "reference list*" OR AB bibliograph* OR AB "hand-search*" OR AB handsearch* OR AB "relevant journals" OR AB "manual search*"
S8	S5 OR S6 OR S7
S9	AB "selection criteria" OR AB "data extraction"
S10	(MH "Literature Review+")
S11	Limiters - Publication Type: Review
S12	S10 OR S11
S13	S9 AND S12
S14	S4 OR S8 OR S13
S15	S14, Limiters - Published Date: 20110101-20210131; Language: English, German
S16	(MH "Internet+") OR (MH "Cellular Phone+")
S17	(MH "Writing+") OR (MH "Reminder Systems")
S18	(MH "Minicomputers")OR (MH "Computers, Hand-Held+")
S19	(MH "Mobile Applications")
S20	TI (mode or modes or method# or way#) N3 (administ* or deliver* or survey# or differen* or collect* or compar* or option# or choice# or conduct* or carr* or complet*) OR AB (mode or modes or method# or way#) N3 (administ* or deliver* or survey# or differen* or collect* or compar* or option# or choice# or conduct* or carr* or complet*)
S21	TI (internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app# or application# or mobile or push or ipad or tablet or handwritten) OR AB (internet or web or online or electronic or paper or pen or postal or pencil or

#	Query
	email or mail* or app# or application# or mobile or push or ipad or tablet or handwritten)
S22	S16 OR S17 OR S18 OR S19 OR S20 OR S21
S23	(MM "Surveys+")
S24	(MM "Questionnaires+")
S25	(MM "Data Collection Methods")
S26	(MM "Outcome Assessment/MT")
S27	(MH "Self Report+")
S28	TI ((questionnaire# or survey#) N3 (internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app# or application# or mobile or push or ipad or tablet or handwritten)) OR AB ((questionnaire# or survey#) N3 (internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app# or application# or mobile or push or ipad or tablet or handwritten))
S29	TI ((patient* or self) N2 report* N2 (outcome* or experien* or measure*)) OR AB ((patient* or self) N2 report* N2 (outcome* or experien* or measure*))
S30	TI (PROM# or PREM# or PRM# or PRO or ePRO#) OR AB (PROM# or PREM# or PRM# or PRO or ePRO#)
S31	S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30
S32	S15 AND S22 AND S31

Tabelle A 5: Suchstrategie für Cochrane via Wiley; Datum der Recherche: 01.02.2021

ID	Search
#1	MeSH descriptor: [Cell Phone] explode all trees
#2	MeSH descriptor: [Internet] explode all trees
#3	MeSH descriptor: [Correspondence as Topic] explode all trees
#4	MeSH descriptor: [Reminder Systems] this term only
#5	MeSH descriptor: [Minicomputers] this term only
#6	MeSH descriptor: [Computers, Handheld] this term only
#7	MeSH descriptor: [Mobile Applications] this term only
#8	MeSH descriptor: [Paper] this term only
#9	((mode or modes or method? or way?) NEAR/3 (administ* or deliver* or survey? or differen* or collect* or compar* or option? or choice? or conduct* or carr* or complet*)) :ti,ab

ID	Search
#10	(internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app? or application? or mobile or push or ipad or tablet or handwritten):ti,ab
#11	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10
#12	[mh "Surveys and Questionnaires"[mj]]
#13	MeSH descriptor: [Data Collection] this term only and with qualifier(s): [methods - MT]
#14	MeSH descriptor: [Outcome Assessment, Health Care] this term only and with qualifier(s): [methods - MT]
#15	MeSH descriptor: [Patient Reported Outcome Measures] this term only
#16	((questionnaire? or survey?) NEAR/3 (internet or web or online or electronic or paper or pen or postal or pencil or email or mail* or app? or application? or mobile or push or ipad or tablet or handwritten)):ti,ab
#17	((patient* or self) NEAR/2 report* NEAR/2 (outcome* or experien* or measure*)):ti,ab
#18	(PROM? or PREM? or PRM? or PRO or ePRO?):ti,ab
#19	#12 or #13 or #14 or #15 or #16 or #17 or #18
#20	#11 and #19 with Cochrane Library publication date Between Jan 2011 and Feb 2021, in Cochrane Reviews

Anhang 2: Anwendung des AMSTAR 2-Instruments im vorliegenden Projekt

Tabelle A 6: Items des AMSTAR-2-Instruments zur Bewertung von systematischen Reviews

Item-Nr.	Wortlaut der originalen AMSTAR-2-Kriterien
1	Did the research questions and inclusion criteria for the review include the components of PICO?
2	Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?
3	Did the review authors explain their selection of the study designs for inclusion in the review?
4	Did the review authors use a comprehensive literature search strategy?
5	Did the review authors perform study selection in duplicate?
6	Did the review authors perform data extraction in duplicate?
7	Did the review authors provide a list of excluded studies and justify the exclusions?
8	Did the review authors describe the included studies in adequate detail?
9	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?
10	Did the review authors report on the sources of funding for the studies included in the review?
11	If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results?
12	If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?
13	Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?
14	Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?
15	If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?
16	Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

Tabelle A 7: Anmerkungen zur projektspezifischen Anwendung von AMSTAR 2

Item	Anmerkungen zur projektspezifischen Anwendung
Item 2	Wenn in einem systematischen Review vermerkt war, dass ein Protokoll angefertigt wurde, wird dies als <i>Partial Yes</i> bewertet. Für eine Bewertung mit <i>Yes</i> muss das Protokoll zusätzlich auffindbar und die Mehrzahl der aufgeführten Unterpunkte erfüllt sein.
Item 4	Der Punkt <i>justified publication restriction</i> unter <i>Partial Yes</i> ist mit einem Punkt unter <i>Yes</i> austauschbar. Insgesamt kann ein <i>Yes</i> vergeben werden, sobald mindestens 2 weitere Punkte unter <i>Yes</i> zutreffen.
Item 8	Sind mindestens 3 Punkte zutreffend, erfolgt die Bewertung mit <i>Partial Yes</i> . Sobald alle Punkte unter <i>Partial Yes</i> zutreffen und eine tabellarische Übersicht der Charakteristika vorhanden ist, wird ein <i>Yes</i> vergeben. Die unter <i>Yes</i> geforderten Punkte erscheinen im Kontext dieses Overviews als zu detailliert bzw. nicht zutreffend, z. B. Dosierung, Follow-up-Zeiträume.
Item 9	Ist mindestens eine Art der Bewertung des Verzerrungsrisikos aufgeführt (nicht zwingend mittels etablierter Checkliste), erfolgt die Bewertung mit <i>Partial Yes</i> . Wurde eine etablierte Checkliste benutzt und sind zudem detailliertere Angaben je Studie (z. B. in einer Tabelle) vorhanden, erfolgt die Bewertung mit <i>Yes</i> .
Item 11	Die Bewertung dieses Items bezieht sich explizit auf die Analyse der in diesem Overview berücksichtigten Vergleiche und Endpunkte. Demnach können andere AMSTAR-2-Bewertungen stark abweichen, da diese sich eventuell ausschließlich auf den primären Endpunkt des zu bewertenden systematischen Reviews beziehen.
Item 14	Das Vorhandensein von Heterogenität wird für diesen Overview bei einem $I^2 \geq 60\%$ definiert. Das Berichten von Ergebnissen verschiedener Effektmodelle bzw. Subgruppen wird als eine Art der Untersuchung bzw. der Diskussion bei vorliegender Heterogenität gewertet.
Item 15	Wird durch die Autoren angegeben, dass bei einer bestimmten Studienanzahl die Bestimmung des Publikationsbias geplant war bzw. nicht durchgeführt werden konnte, ist das Item mit <i>Yes</i> zu bewerten.
Item 16	Bei einer Angabe von potenziellen Interessenkonflikten inkl. einer Angabe zur Finanzierung erfolgt die Bewertung mit <i>Yes</i> . Eine Beschreibung des Managements bei vorliegenden Interessenkonflikten ist im Rahmen dieses Overviews nicht erforderlich.

Anhang 3: Ausgeschlossene Artikel in der systematischen Literaturrecherche**Nicht E1**

1. Brewer, SE; Morrato, EH (2016): 546. Regulatory Considerations for Knowledge, Attitude, and Belief Surveys: A Systematic Review of Physician Survey Response [Abstract]. *Pharmacoepidemiology and Drug Safety* 25(Suppl. 3): 319. DOI: 10.1002/pds.4070.
2. Brueton, V; Tierney, J; Stenning, S; Nazareth, I; Meredith, S; Harding, S; et al. (2013): A23: Systematic Review of Strategies to Reduce Attrition in Randomised Trials [Abstract]. *Clinical Trials* 10: S11. DOI: 10.1177/1740774513497438.
3. Fluss, E; Macfarlane, GJ; Bond, C; Jones, GT (2012): 297. Do Electronic and Paper Questionnaires Produce Equivalent Results? A Systematic Review [Poster Viewing III. Basic science]. *Rheumatology* 51: iii163-iii164. DOI: 10.1093/rheumatology/kes108.
4. Gore, S; Blackwood, J; Chindam, T; Goldberg, A; Huang, MH; Shoemaker, MJ; et al. (2018): Validity and Reliability of Subjective Physical Activity Assessments in COPD: A Systematic Review [Abstract]. *Cardiopulmonary Physical Therapy Journal* 29(1): 26-26. DOI: 10.1097/CPT.000000000000077.
5. White, MK; Bjorner, JB (2016): (2052) A review of measurement equivalence study findings of the SF-36 Health Survey across computer, tablet, and handheld devices compared to paper administration [Abstract]. *Quality of Life Research* 25: 143. DOI: 10.1007/s11136-016-1390-7.
6. Wild, D; Muehlhausen, W; O Donohoe, P; Doll, H (2014): PRM100: Equivalence of Electronic and Paper Administration of Patient Reported Outcome Measures: A Systematic Review and Metaanalysis. *Value in Health* 17(3): A198. DOI: 10.1016/j.jval.2014.03.1156.

Nicht E4

1. Brueton, VC; Tierney, JF; Stenning, S; Meredith, S; Harding, S; Nazareth, I; et al. (2014): Strategies to improve retention in randomised trials: a Cochrane systematic review and meta-analysis. *BMJ open* 4: e003821. DOI: 10.1136/bmjopen-2013-003821.

Nicht E5

1. Ahmad, I; Ali, PA; Rehman, S; Talpur, A; Dhingra, K (2017): Intimate partner violence screening in emergency department: a rapid review of the literature. *Journal of Clinical Nursing* 26(21-22): 3271-3285. DOI: 10.1111/jocn.13706.
2. Angermeyer, MC; Schomerus, G (2017): State of the art of population-based attitude research on mental health: a systematic review. *Epidemiology and Psychiatric Sciences* 26(3): 252-264. DOI: 10.1017/S2045796016000627.

3. Bahkali, S; Almainan, A; Al-Nasser, L; Elmetwally, A; Househ, M (2014): Are Web-based Surveys the New Epidemiological Mode For Healthcare Research? - The Saudi Perspective. *Integrating Information Technology and Management for Quality of Care* 202: 189-192. DOI: 10.3233/978-1-61499-423-7-189.
4. Basílio, N; Cardoso, S; Nunes, JM; Laranjo, L; da Luz Antunes, M; Heleno, B (2018): Portuguese Primary Care physicians response rate in surveys: A systematic review. *Revista da Associação Médica Brasileira* 64(3): 272-280. DOI: 10.1590/1806-9282.64.03.272.
5. Byrom, B; Gwaltney, C; Slagle, A; Gnanasakthy, A; Muehlhausen, W (2019): Measurement Equivalence of Patient-Reported Outcome Measures Migrated to Electronic Formats: A Review of Evidence and Recommendations for Clinical Trials and Bring Your Own Device. *Therapeutic Innovation & Regulatory Science* 53(4): 426-430. DOI: 10.1177/2168479018793369.
6. Corsini, N; Fish, J; Ramsey, I; Sharplin, G; Flight, I; Damarell, R; et al. (2017): Cancer survivorship monitoring systems for the collection of patient-reported outcomes: a systematic narrative review of international approaches. *Journal of Cancer Survivorship* 11(4): 486-497. DOI: 10.1007/s11764-017-0607-2.
7. Ekhtiari, S; Kay, J; de Sa, D; Simunovic, N; Musahl, V; Peterson, DC; et al. (2017): What Makes a Successful Survey? A Systematic Review of Surveys Used in Anterior Cruciate Ligament Reconstruction. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 33(5): 1072-1079.e3. DOI: 10.1016/j.arthro.2017.01.032.
8. Khanbhai, M; Flott, K; Darzi, A; Mayer, E (2019): Evaluating Digital Maturity and Patient Acceptability of Real-Time Patient Experience Feedback Systems: Systematic Review. *Journal of Medical Internet Research* 21(1): e9076. DOI: 10.2196/jmir.9076.
9. Kilsdonk, E; Wendel, E; van Dulmen-den Broeder, E; van Leeuwen, FE; van den Berg, MH; Jaspers, MW (2017): Participation rates of childhood cancer survivors to self-administered questionnaires: a systematic review. *European Journal of Cancer Care* 26(6): e12462. DOI: 10.1111/ecc.12462.
10. McCluskey, S; Topping, AE (2011): Increasing response rates to lifestyle surveys: A pragmatic evidence review. *Perspectives in Public Health* 131(2): 89-94. DOI: 10.1177/1757913910389423.
11. Meyer, VM; Benjamens, S; Mourni, ME; Lange, JFM; Pol, RA (2020): Global Overview of Response Rates in Patient and Health Care Professional Surveys in Surgery: A Systematic Review. *Annals of Surgery* Online ahead of print. DOI: 10.1097/sla.0000000000004078.
12. Pierannunzi, C; Hu, SS; Balluz, L (2013): A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004-2011. *BMC: Medical Research Methodology* 13:49. DOI: 10.1186/1471-2288-13-49.
13. Şahin, F; Yan, Z (2013): Mobile Phones in Data Collection: A Systematic Review. *International Journal of Cyber Behavior, Psychology and Learning* 3(3): 67-87. DOI: 10.4018/ijcbpl.2013070106.

14. Tran, TXM; Park, J; Lee, J; Jung, Y-S; Chang, Y; Cho, H (2021): Utility of the Patient-Reported Outcomes Measurement Information System (PROMIS) to measure primary health outcomes in cancer patients: a systematic review. *Supportive Care in Cancer* 29: 1723–1739. DOI: 10.1007/s00520-020-05801-6.
15. Wang, K; Eftang, CN; Jakobsen, RB; Aroen, A (2020): Review of response rates over time in registry-based studies using patient-reported outcome measures. *BMJ Open* 10(8): e030808. DOI: 10.1136/bmjopen-2019-030808.
16. Yarger, JB; James, TA; Ashikaga, T; Hayanga, AJ; Takyi, V; Lum, Y; et al. (2013): Characteristics in response rates for surveys administered to surgery residents. *Surgery* 154: 38-45. DOI: 10.1016/j.surg.2013.04.060.
17. Zha, N; Alabousi, M; Katz, DS; Su, J; Patlas, M (2020): Factors Affecting Response Rates in Medical Imaging Survey Studies. *Academic Radiology* 27(3): 421-427. DOI: 10.1016/j.acra.2019.06.005.

Nicht E7

1. Bernstein, DN; McIntyre, AW; Baumhauer, JF (2020): Effect of assessment administration method and timing on patient-reported outcome measures completion and scores: Overview and recommendations. *Musculoskeletal Care* 18(4): 535-540. DOI: 10.1002/msc.1476.
2. Cho, Yi; Johnson, TP; Vangeest, JB (2013): Enhancing Surveys of Health Care Professionals: A Meta-Analysis of Techniques to Improve Response. *Evaluation & the Health Professions* 36(3): 382-407. DOI: 10.1177/0163278713496425.
3. Cornesse, C; Bosnjak, M (2018): Is there an association between survey characteristics and representativeness? A meta-analysis. *Survey Research Methods* 12(1): 1-13. DOI: 10.18148/srm/2018.v12i1.7205.
4. Daikeler, J; Bošnjak, M; Lozar Manfreda, K (2019): Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. *Journal of Survey Statistics and Methodology* 8(3): 513-539. DOI: 10.1093/jssam/smz008.
5. Gnambs, T; Kaspar, K (2015): Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis. *Behavior Research Methods* 47(4): 1237-1259. DOI: 10.3758/s13428-014-0533-4.
6. Martins, Y; Lederman, RI; Lowenstein, CL; Joffe, S; Neville, BA; Hastings, BT; et al. (2012): Increasing response rates from physicians in oncology research: a structured literature review and data from a recent physician survey. *British Journal of Cancer* 106: 1021-1026. DOI: 10.1038/bjc.2012.28.
7. 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. DOI: 10.1093/poq/nfs047.

8. Pit, SW; Vo, T; Pyakurel, S (2014): The effectiveness of recruitment strategies on general practitioner's survey response rates – a systematic review. *BMC: Medical Research Methodology* 14:76. DOI: 10.1186/1471-2288-14-76.
9. Pombo, N; Bousson, K; Araujo, P; Viana, J (2015): Medical decision-making inspired from aerospace multisensor data fusion concepts. *Informatics for Health and Social Care* 40(3): 185-197. DOI: 10.3109/17538157.2013.872113.
10. VanGeest, J; Johnson, TP (2011): Surveying Nurses: Identifying Strategies to Improve Participation. *Evaluation & the Health Professions* 34(4): 487-511. DOI: 10.1177/0163278711399572.

Nicht E8

1. Brueton, VC; Tierney, J; Stenning, S; Harding, S; Meredith, S; Nazareth, I; et al. (2013): Strategies to improve retention in randomised trials (Review). *Cochrane Database of Systematic Reviews* (12): Art. No.: MR000032. DOI: 10.1002/14651858.MR000032.pub2.
2. Clifford, S; Perez-Nieves, M; Skalicky, AM; Reaney, M; Coyne, KS (2014): A systematic literature review of methodologies used to assess medication adherence in patients with diabetes. *Current Medical Research and Opinion* 30(6): 1071-1085. DOI: 10.1185/03007995.2014.884491.
3. Hood, K; Robling, M; Ingledew, D; Gillespie, D; Greene, G; Ivins, R; et al. (2012): Mode of data elicitation, acquisition and response to surveys: a systematic review. *Health Technology Assessment* 16(27): 1-162. DOI: 10.3310/hta16270.
4. Remillard, ML; Mazor, KM; Cutrona, SL; Gurwitz, JH; Tjia, J (2014): Systematic Review of the Use of Online Questionnaires of Older Adults. *Journal of the American Geriatrics Society* 62(4): 696-705. DOI: 10.1111/jgs.12747.
5. Smiley, SL; Milburn, NG; Nyhan, K; Taggart, T (2020): A Systematic Review of Recent Methodological Approaches for Using Ecological Momentary Assessment to Examine Outcomes in U.S. Based HIV Research. *Current HIV/AIDS Reports* 17(4): 333-342. DOI: 10.1007/s11904-020-00507-0.
6. van Gelder, MMHJ; Vlenterie, R; IntHout, J; Engelen, LJLPG; Vrieling, A; van de Belt, TH (2018): Most response-inducing strategies do not increase participation in observational studies: a systematic review and meta-analysis. *Journal of Clinical Epidemiology* 99: 1-13. DOI: 10.1016/j.jclinepi.2018.02.019.
7. VanGeest, JB; Johnson, TP (2012): Using incentives in surveys of cancer patients: do "best practices" apply? *Cancer Causes & Control* 23(12): 2047-2052. DOI: 10.1007/s10552-012-0082-z.
8. Warrington, L; Absolom, K; Conner, M; Kellar, I; Clayton, B; Ayres, M; et al. (2019): Electronic Systems for Patients to Report and Manage Side Effects of Cancer Treatment: Systematic Review. *Journal of Medical Internet Research* 21(1): e10875. DOI: 10.2196/10875.

Anhang 4: Charakteristika der eingeschlossenen systematischen Reviews

Tabelle A 8: Charakteristika der eingeschlossenen systematischen Reviews

Erstautor/in (Jahr)	Datenbanken, andere Quellen	Suchzeitraum	Population ^{*3}	Alter (MW ± SD)	Fragebögen/ Instrumente	eingeschlossene Studien	Studiendesigns	Col	Finanzierung
Alfonsson et al. (2014) ^{*1}	Embase, MEDLINE, PsycINFO; reference lists	up to Jan 2014	mixed	18.8 to 68.3	self-report instruments designed to measure symptoms of selected DSM-IV Axis I group diagnoses	38	9 RCT, 6 non-RCT, 18 crossover, 5 two samples	no	n.s.
Blumenberg und Barros (2018) ^{*2}	LILACS, PubMed, SciELO, WebSM; Google Scholar; reference lists	up to Mar 2017	healthy people	n.s.	surveys had to be about health-related topics	19	9 RCT, 10 cross-sectional	no	no
Campbell et al. (2015) ^{*1}	BIOSIS Citation Index, Embase, MEDLINE, PsycINFO, PubMed, SciELO Citation Index, Web of Science Core Collection	2007 to Mar 2014	mixed	n.s.	validated PRO instruments	55	2 RCT, 29 rand. crossover, 17 crossover, 4 comparison, 3 unclear	no	n.s.
Dodou und de Winter (2014) ^{*1}	Web of Knowledge; Google Scholar; reference list	up to Jan 28, 2014	mixed	n.s.	social desirability or lie scales	51	between and within study (rand. and non-rand.)	n.s.	n.s.

Erstautor/in (Jahr)	Datenbanken, andere Quellen	Suchzeitraum	Population ^{*3}	Alter (MW ± SD)	Fragebögen/ Instrumente	eingeschlossene Studien	Studiendesigns	Col	Finanzierung
Elfeky et al. (2020) ^{*2}	CINAHL, Cochrane Controlled Trials Register, Cochrane Methodology Register, MEDLINE	2007 to Oct 2017	patients	n.s.	<i>follow-up questionnaires in clinical trials (no further details)*4</i>	14	non-RCT	no	no
Flanagan et al. (2015) ^{*1,2}	Assia, Embase, CINAHL, Cochrane, CLib, ERIC, MEDLINE, MEDLINE in Process, SSCI	up to Feb 2013	mixed	8 to 32 (n = 20 studies)	<i>health related data collection methods in children, teenagers and young people</i>	30	<i>comparison</i>	no	no
Gnambs und Kaspar (2017) ^{*1}	Business Source Complete, EconLit, ProQuest Dissertations & Theses Database, PsycINFO, Psychology & Behavioral Sciences Collection, Psycindex; Google Scholar	n.s. <i>(studies of the current millennium)</i>	n.s.	22.42 ± 5.70, 23.88 ± 6.59, 32.51 ± 11.25 ^{*5}	validated social desirability scales	47	<i>RCT, crossover</i>	no	no
Greenleaf et al. (2017) ^{*1,2}	7 databases	n.s.	n.s.	n.s.	survey includes questions about health	10 ^{*6}	1 RCT, 8 cross-sectional, 1 unclear	no	no
Hussain et al. (2015) ^{*1}	Embase, CINAHL, Cochrane Library, DARE, MEDLINE, PsycINFO	up to Jan 2011	mixed	≥18	screening for intimate partner violence	6	RCT	no	no
Jibb et al. (2020) ^{*1,2}	Embase, CENTRAL, MEDLINE	up to Nov 19, 2019	patients	41.5 ± 17.5 (n	acute or chronic pain assessments	74	8 RCT, 19 rand. crossover, 5	no	n.s.

Erstau- tor/in (Jahr)	Datenbanken, andere Quel- len	Such- zeit- raum	Popula- tion ^{*3}	Alter (MW ± SD)	Fragebögen/ Instrumente	einges- chlos- sene Stu- dien	Studiendesigns	Col	Finan- zierung
				= 55 studies)			non-rand. cross- over, 5 crosso- ver (rand. un- clear), 1 rand. cohort, 14 non- rand. cohort, 1 unclear		
Marcano Belisario et al. (2015) ^{*1}	ACM Digital, CAB Abstracts, Campbell Library, CENTRAL, Currents Contents Connect, Embase, ERIC, Health Man- agement Information Con- sortium, IEEEExplore, MED- LINE, PsycINFO, Sociological Abstracts, Web of Science; clinical trial registers; grey literature; reference lists; Google Scholar	2007 to Apr 13, 2015	mixed	4 to 11 and 18 to 80	health-related self-admin- istered survey question- naires (validated and non- validated)	14	RCT, crossover, paired repeated measure study	no	no
Meirte et al. (2020) ^{*1}	PubMed, Web of Science; reference lists	up to Oct 2017	mixed	49	PROMs	32	experimental (crossover or parallel), obser- vational	no	n.s.

Erstautor/in (Jahr)	Datenbanken, andere Quellen	Suchzeitraum	Population ^{*3}	Alter (MW ± SD)	Fragebögen/ Instrumente	eingeschlossene Studien	Studiendesigns	Col	Finanzierung
Muehlhausen et al. (2015) ^{*1}	Embase, MEDLINE, PsycINFO; reference lists; grey literature; clinical trials	Jan 2007 to Dec 2013	mixed	42.9 ± 17.1	PROMs	72	RCT and non-RCT (parallel groups or cross-over)	no	n.s.
Rutherford et al. (2016) ^{*1,2}	Embase, CINAHL, MEDLINE, PsycINFO, Scopus; reference lists; author search	2004 to Apr 2014	mixed	n.s.	patient-reported outcome measures intended for assessing health outcomes including quality of life, symptoms, psychological aspect, satisfaction	56	31 RCT, 12 cross-sectional, 3 longitudinal, 10 unclear	no	no
Weigold et al. (2018a) ^{*1}	Business Source Complete, MEDLINE, PsycINFO; reference lists	up to May 2013	mixed	n.s.	structured questionnaires that participants completed about themselves on a wide range of topics such as attitudes, behaviors, and personal characteristics	159	between and within study	n.s.	n.s.
White et al. (2018) ^{*1}	Embase, MEDLINE, PubMed, Reuters; Optum database; reference lists	up to Oct 2016	mixed	≥ 18	SF-12 (v1 or v2), SF-36 (v1 or v2)	25	2 RCT, 14 rand. crossover, 9 non-rand. cross-over	yes	yes

^{*1} outcome data quality, ^{*2} outcome response rate, ^{*3} mixed: healthy people and patients, ^{*4} study on participants retention in RCTs, ^{*5} paper on 3 meta-analysis, probably 53 % of participants were proctored, ^{*6} only studies from low- and middle-income countries

Abkürzungen: CoI = potential conflict of interest; DSM IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; MW = Mittelwert; n.s. = not stated; PRO = patient-reported outcome; RCT = randomized controlled trial; SD = Standardabweichung;
kursiv: eigene Berechnung/Interpretation

Anhang 5: AMSTAR-2-Bewertung der eingeschlossenen systematischen Reviews

Tabelle A 9: AMSTAR-2-Bewertung der eingeschlossenen systematischen Reviews

Referenz	AMSTAR 2-Kriterium																	
	1	2	3	4	5	6	7	8	9		10	11		12	13	14	15	16
									RCT	NRSI		RCT	NRSI					
Alfonsson et al. (2014)	Y	N	N	PY	Y	Y	N	Y	PY	PY	N	MA	MA	M A	N	N	M A	N
Blumenberg und Barros (2018)	Y	N	N	PY	N	N	N	PY	N	N	N	MA	MA	M A	N	Y	M A	Y
Campbell et al. (2015)	Y	N	N	PY	N	N	N	PY	N	N	N	MA	MA	M A	N	N	M A	N
Dodou und de Winter (2014)	Y	N	N	N	N	N	N	PY	N	N	N	Y	Y	N	N	N	Y	N
Elfeky et al. (2020)	Y	Y	Y	Y	Y	Y	N	Y	NRSI	Y	N	MA	MA	M A	Y	Y	M A	Y
Flanagan et al. (2015)	Y	N	N	PY	N	N	N	N	N	N	N	MA	MA	M A	N	N	M A	Y
Gnambs und Kaspar (2017)	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	Y	Y

Referenz	AMSTAR 2-Kriterium																	
	1	2	3	4	5	6	7	8	9		10	11		12	13	14	15	16
									RCT	NRSI			RCT	NRSI				
Greenleaf et al. (2017)	Y	N	N	N	N	Y	N	N	N	N	N	MA	MA	MA	N	N	MA	Y
Hussain et al. (2015)	Y	N	N	PY	Y	Y	N	Y	Y	RCT	N	Y	MA	N	N	Y	N	Y
Jibb et al. (2020)	Y	PY	N	PY	Y	Y	N	Y	N	N	N	MA	MA	MA	N	Y	MA	N
Marcano Belisario et al. (2015)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	MA	MA	MA	Y	Y	MA	Y
Meirte et al. (2020)	Y	PY	N	PY	N	N	N	Y	PY	PY	N	MA	MA	MA	Y	N	MA	N
Muehlhausen et al. (2015)	Y	N	N	Y	N	N	N	Y	N	N	N	N	N	N	N	Y	Y	N
Rutherford et al. (2016)	Y	N	N	PY	Y	Y	N	PY	PY	PY	N	N	N	N	N	Y	N	Y
Weigold et al. (2018a)	Y	N	N	PY	N	Y	N	N	N	N	N	N	N	N	N	Y	N	N
White et al. (2018)	Y	N	N	PY	N	N	N	Y	N	N	N	N	N	N	N	N	N	Y

Abkürzungen: Y = Yes; PY= partial yes; N = No; NRSI = includes only NRSI; RCT = includes only RCT; MA = no meta-analysis conducted

farbig: critical domains

Anhang 6: Ergebnisse der systematischen Literaturrecherche: Response-Rate

Tabelle A 10: Ergebnisse zur Response-Rate

Erstautor/in (Jahr)	Eingeschlossene Studien	Analytisierte Fälle	Population analysiert	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
1 Web/computer versus paper								
Blumenberg und Barros (2018)	6 RCT, 3 cross-sectional	77.952	n.s.	questionnaire located in a web page accessible using a web-browser	letters had to be sent via traditional mail (postal)	response rate (according to standard American Association for Public Opinion Research definition)	web: 51.9 %; mail: 59.7 % web: 50.3 %; mail: 64.4 % web: 50.9 %; mail: 64.4 % web: 15.4 %; mail: 43.6 % web: 26.3 %; mail: 53.5 % web: 93.4 %; mail: 89.1 % web: 67.0 %; mail: 72.9 % web: 53.4 %; mail: 56.1 % web: 53.3 %; mail: 56.1 % web: 28.7 %; mail: 49.8 % web: 83.8 %; mail: 80.6 %	included two 3-arm studies
Elfeky et al. (2020)	1 cohort	1.329	young female students	online questionnaire	postal questionnaire	4 weeks and 3 months after follow-up commenced	4 weeks: web: 51 %; mail: 29 %; RR 1.78 [1.47, 2.14]; 3 months: web: 72 %; mail: 59 %	study on participants retention in RCTs
Flanagan et al. (2015)	2 unclear	≥1.323	12 - 16 and 18 - 28 years	web-based computerized	paper	response rate	web-based mode had a better response rate than paper	-

Erstau- tor/in (Jahr)	Einge- schlossene Studien	Analy- sierte Fälle	Population analy- siert	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
Jibb et al. (2020)	2 RCT	2.797	diabetes, asthma, heart disease, lung disease, hyperten- sion, total hip re- placement surgical pain	computer	paper-based tool	percent of participants who com- pleted as- sessments	computer: 87.5 %, paper: 83.1 % (P=.19), computer: 49 %, paper: 92 % (P<.01)	-
Ruther- ford et al. (2016)	1 cross-sec- tional study (nonre- sponse: al- ternative mode of- fered), 8 <i>unclear</i>	>800 (n = 1 study)	benign abnormality or breast cancer (n = 1 study)	electronic (e. g. compu- ter)	paper (e. g. at home/ postal)	response rate	higher response rates for paper (range 61 to 89 %) compared to electronic (range 18 to 39 %, 2 stud- ies); favor of electronic (39 to 90 %) compared to paper (22 to 73 %, 3 stud- ies); no difference (4 stud- ies)	study and pa- tients charac- teristics re- ported for 1 out of 9 stud- ies only

2 Mobile phone versus paper

Jibb et al. (2020)	1 RCT	40	posthysterectomy and postcholecys- tectomy pain	mobile phone	paper-based tool	percent of participants completing assessments	day of surgery: mobile phone: 35 %, paper: 41 %; after 2 - 4 days: mobile phone: 100 %, paper: 100 %	-
-----------------------	-------	----	---	-----------------	---------------------	---	--	---

Abkürzungen: n.s. = not stated; RCT = randomized controlled trial; RR = risk ratio

kursiv: eigene Berechnung/Interpretation

Anhang 7: Ergebnisse der systematischen Literaturrecherche: Antwortverhalten – Vergleiche mit Paper-Pencil-Befragung

Tabelle A 11: Ergebnisse zur Übereinstimmung des Antwortverhaltens – Vergleiche mit Paper-Pencil-Befragung

Erstautor/in (Jahr)	Eingeschlossene Studien	Analy-sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
1 Electronic vs paper							
Campbell et al. (2015)	29 rand. crossover, 2 rand. comparison, 17 crossover, 4 comparison, 3 unclear	>6.602 (11 studies stated 'multiple')	validated electronic PRO instrument* ¹	PnP PRO instrument (paper-based original PRO instrument)	measurement equivalence (author conclusion based on 44 correlation coefficients [ICC, pearson's r, spearman's rho, kappa coefficient, unclear, other])	equivalence: 43 studies; failed to find equivalence: 2 studies; unclear result: 10 studies (e. g. finding equivalence for some instruments but not for all)	-
Rutherford et al. (2016)	4 RCT, 20 rand. crossover, 1 comparison, 6 cross-sectional	13.061	paper (hard copy at home/postal or in clinic)	electronic	difference in mean scores	0.01 95 % CI [-0.02, 0.04]	included 1 study investigating CATI and 2 non-controlled studies
Weigold et al. (2018a)	111 between-group	103.445	paper	computer* ²	equivalence (mean difference, Hedges' g)	-0.024 95 % CI [-0.065, 0.017]	I ² = 89.23 %
	52 within-subject	≥5.045 participants	paper	computer* ²	equivalence (mean difference, Hedges' g)	-0.020 95 % CI [-0.038, 0.002]	I ² = 17.61 %
White et al. (2018)	11 rand. crossover, 9 non-rand. crossover	3.444	paper	electronic (PC [8], web [3], tablet [5],	agreement (ICC [n = 12], Pearson's r [n = 7], Spearman's rho [n = 1])	RE-scale: 0.76; RP: 0.83; PF: 0.91 [CI not reported]; for all subscales	results for other subscales shown in Fig. 2b only (no

Erstau- tor/in (Jahr)	Eingeschlossene Studien	Analy- sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
				handheld [1], PDA [2], tab- let/web [1])* ³		the 95 % CI was above 0.7	numerical data re- ported)
	2 RCT, 10 rand. crossover, 3 non- rand. crossover	3.881	paper	electronic (PC [8], web [3], tablet [3], PDA [1])* ³	difference in means (con- verted to norm-based scores)	PF: 0.01; GH: - 0.39 [CI <i>not reported</i>]; all 95 % CI were within the speci- fied ± 2 point margin of equivalence	results for other subscales shown in Fig. 2a only (no numerical data re- ported)
2 Web/computer vs paper							
Alfons- son et al. (2014)	1 RCT, 2 non-RCT, 1 crossover	563	pen and paper	online	interformat reliability (ICC)	0.54 to 0.96	-
	2 non-RCT, 7 crossover	1.624	pen and paper	online	interformat reliability (Pearson's r)	0.30 to 0.99	-
	5 RCT, 2 crossover	4.600	pen and paper	online	interformat reliability (mean difference)	not significant (6 stud- ies), significant (1 study)	-
	2 non-RCT, 4 crossover	814	pen and paper	computer	interformat reliability (Pearson's r)	0.35 to 0.98	-
	3 RCT, 2 crossover	862	pen and paper	computer	interformat reliability (mean difference)	not significant (4 stud- ies), significant (1 study)	-
Dodou und de	51 between and within study (rand. and non-rand.)	16.700 <i>partici- pants</i>	computer (computer- paper)	paper-and- pencil	social desirability (SMD)	FE: 0.00 95 % CI [-0.03, 0.03]; RE: -0.01 95 % CI [-0.07, 0.04]; subgroup	I ² = 63.5 %; prob- ably one paper

Erstautor/in (Jahr)	Eingeschlossene Studien	Analy-sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
Winter (2014)			ized surveys, online or offline)			analyses: 5 and 4 out of 30 effect sizes for FE and RE were significantly different (FE: sensitive nature of question, BIDR-SD, MCDS, randomised/ within-subjects and non-randomised assignment; RE: BIDR-IM, BIDR-SD, MCDS, non-randomised assignment)	falsely cited as included study (Wicherts 2012); fulltexts for 3 studies could not be obtained by authors, data were gathered from a previous meta-analysis
Flanagan et al. (2015)	2 unclear	≥1.238	web-based computerized	PnP	reporting of risk behaviour and substance use	higher reported rates in paper	-
	2 unclear	≥1.759	web-based computerized	PnP	incompleteness of items	higher level of incompleteness in paper; in web-based mode respondents were more likely to skip questions (1 study each)	-
Gnambs und	12 unclear	3.746 participants	web-based surveying and testing	paper-and-pencil	social desirability distortion (SMD [adjusted effect])	0.03 90 % credibility interval [-0.03, 0.08]	probably 53 % of participants were proctored

Erstautor/in (Jahr)	Eingeschlossene Studien	Analy-sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
Kaspar (2017)	10 unclear	2.951 participants	web-based surveying and testing	paper-and-pencil	Big Five of personality (SMD [adjusted effect])	0.05 90 % credibility interval [-0.12, 0.21]	probably 79 % of participants were proctored
	28 unclear	16.034 participants	web-based surveying and testing	paper-and-pencil	psychopathology (SMD [adjusted effect])	0.00 90 % credibility interval [-0.11, 0.11]	probably 54 % of participants were proctored
Hussain et al. (2015)	2 RCT	2.105	computer-assisted screen	self-administered written screen	rate of intimate partner violence disclosure	OR: 1.23 95 % CI [0.92, 1.64]	-
Jibb et al. (2020)	2 RCT, 5 rand. crossover, 1 non-rand. crossover, 1 crossover (rand. unclear)	4.057	computer	paper	pain score equivalence (e. g. ICC, spearman rho, t test, Mann-Whitney U test)	equivalence (6 studies), nonequivalence (1 study), discrepant results (2 studies)	results interpreted by review authors
Meirte et al. (2020)	2 between study design (rand. or non-rand.), 2 rand. crossover	≥1.760 (n = 3 studies)	PC (web form)	paper	data completion, missing data, completion rate (several definitions)	not extractable due to different and inconsistent definitions of outcomes	-
	1 between study (rand. unclear)	126	web	paper	data completion, data missing (unclear definitions)	not extractable due to different and inconsistent definitions of outcomes	-

Erstau- tor/in (Jahr)	Eingeschlossene Studien	Analy- sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
	3 rand. crossover	>468 (n = 2 studies)	PC (web, PC)	paper	agreement (ICC, r, CCC)	0.86 to 1.00 (qualitative variables); 0.69 to 1.00 for 18 qualitative variables (height, weight, hip circumference, waist circumference were all different); 5 SAQ domains r = 0.84 to 0.93; 8 SF-36 subscales: r = 0.54 to 0.75; >0.849 [CI not reported] (p > 0.05)	-
	1 rand. crossover	80	PC	paper	agreement (ICC)	0.982 [CI not reported]	-
Muehl- hausen et al. (2015)	20 rand. crossover, 6 rand. 3/4 group crossover, 1 non-rand. crossover, 1 non-rand. 3/4 group crossover, 6 within subject (non-rand.)	>5.718 (n = 26)	paper	web/PC	correlation	RE: 0.87 95 % CI [0.85, 0.89]	I ² = 92.2 %
	5 RCT, 2 non-RCT, 3 rand. crossover, 1 within subject (non-rand.)	8.902	paper	web/PC	equivalence (SMD)	0.006 to 0.052	-

Erstau- tor/in (Jahr)	Eingeschlossene Studien	Analy- sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
3 Smartphone vs paper							
Alfonson et al. (2014)	1 RCT	45 participants in 3 arms	pen and paper	smartphone	interformat reliability (ICC)	0.90 to 0.92	-
Jibb et al. (2020)	1 RCT, 4 rand. crossover, 1 non-rand. crossover, 1 crossover (rand. unclear), 2 non-rand. cohort	1.414	mobile phone	paper (verbally administered tool, paper-based tool)	pain score equivalence (e. g. pearson correlation, ICC)	equivalence (8 studies), discrepant results (1 study)	results interpreted by review authors
Marcano Belisario et al. (2015)	1 RCT, 4 crossover	3.596	app	paper	data equivalence (mean score differences in validated survey questionnaires)	no statistically significant difference	-
	5 crossover	3.562	app	paper	data equivalence (ICC, correlation coefficient)	4 studies exceeded threshold of 0.70; 1 study exceeded threshold of 0.60	0.60 interpreted as 'equivalent' by review author
	1 crossover	80	app	paper	data equivalence (mean score differences in non-validated survey questionnaires)	no statistically significant difference	-

Erstau- tor/in (Jahr)	Eingeschlossene Studien	Analy- sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
	1 crossover	80	app	paper	data equivalence (correla- tion coefficient r)	exceeded treshold of 0.60	-
	1 crossover	42	app	paper	data completeness (mean number of complete rec- ords)	MD: 7.08 95 % CI [2.90, 11.26]	-
	1 crossover	42	app	paper	number in incomplete records	fewer incomplete rec- ords when using the app	-

4 Tablet vs paper

Jibb et al. (2020)	2 rand. crossover, 1 non-rand. crosso- ver, 1 non-rand. co- hort	558	tablet	paper	pain score equivalence (ICC, pearson correlation, spearman rho, t test)	equivalence (3 studies), discrepant results (1 study)	results inter- preted by review authors
Meirte et al. (2020)	2 between study (rand. unclear), 2 rand. crossover	>1.760 (n = 3 studies)	tablet (web form)	paper	data completion, missing data, completion rate (several definitions)	<i>not extractable due to different and incon- sistent definition of out- comes</i>	-
	4 rand. crossover	1.240	tablet (web, PC)	paper	agreement (ICC, r)	ICC: 0.98 95 % CI [0.97, 0.99]; ICC: >0.9 [CI not reported] (p <0.0001); r: 0.87 to 0.98 (p >0.05); no significant differences	-

Erstau- tor/in (Jahr)	Eingeschlossene Studien	Analy- sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
Muehl- hausen et al. (2015)	6 rand. crossover, 1 within subject (non-rand.)	1.952	paper	tablet (touch)	correlation	RE: 0.90 95 % CI [0.86, 0.94]	I ² = 96.6 %; ana- lyzed samples in- cluded 1 3-arm study (n = 61)
	1 within subject (rand.)	112	paper	tablet (touch)	equivalence (SMD)	0.035	-

5 Palm/PDA vs paper

Alfons- son et al. (2014)	1 crossover	160	pen and paper	palm	interformat reliability (ICC)	0.99	-
	1 crossover	1.512	pen and paper	PDA	interformat reliability (mean difference)	not significant	-
Jibb et al. (2020)	2 RCT, 5 rand. crossover, 1 crosso- ver (rand. unclear), 2 non-rand. crosso- ver, 4 non-rand. co- hort	2.286	PDA	paper	pain score equivalence (t test, pearson correlation, Mann-Whitney test, wheighted kappa, spearman rho, ICC, ANOVA, Wilcoxon' signed rank test)	equivalence (12 studies), discrepant results (2 studies)	results inter- preted by review authors
Meirte et al. (2020)	1 rand. crossover	296	PC (PDA)	paper	agreement (ICC)	0.92 [CI not reported]	-
	1 rand. crossover	232	web (PDA)	paper	agreement	AQLQ: p = 0.009, ACQ: p = 0.12, RQLQ: p = 0.05	interpretation un- clear (to less infor- mation reported)

Erstau- tor/in (Jahr)	Eingeschlossene Studien	Analy- sierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
Muehl- hausen et al. (2015)	6 rand. crossover, 1 rand. 3/4 group crossover, 3 within subject (non-rand.)	≥ 1.504 (n = 9 studies)	paper	PDA	correlation	RE: 0.88 95 % CI [0.84, 0.92]	I ² = 77.7 %
	3 rand. crossover, 1 non-rand. crosso- ver	2.060	paper	PDA	equivalence (SMD)	0.010 to 0.026	-

6 IVRS vs paper

Green- leaf et al. (2017)	1 RCT	1.810	in-person IVR	postal SAQ* ⁴	rate of socially desirable answers	more frequently re- ported in postal SAQ	aim: studies from low- and middle- income countries
	1 RCT	1.810	in-person IVR	postal SAQ* ⁴	item nonresponse rate	statistically significantly different (<i>no further de- tails</i>)	aim: studies from low- and middle- income countries
	1 RCT	1.860	in-person IVR	in-person SAQ (on-site)* ⁵	item nonresponse rate	statistically significantly different (<i>no further de- tails</i>)	aim: studies from low- and middle- income countries
Muehl- hausen et al. (2015)	4 rand. crossover, 1 rand. 3/4 group crossover, 3 within subject (non-rand.)	>1.354 (n = 7 studies)	paper	IVRS	correlation	RE: 0.87 95 % CI [0.83, 0.91]	I ² = 81.2 %

*¹ including internet, tablet, computer, touch-screen computer, PDA; *² including most electronic devices on which participants could manually complete surveys, excluding devices into which participants spoke their responses and those that employed computer adaptive testing; *³ tablet: ≥ 7 inch diameter screen, handheld: e. g. smartphone or PDA with < 7 inch diameter screen; *⁴ completed off-site, returned via post; *⁵ completed on-site, returned in a self-sealing envelope

Abkürzungen: CATI = computer-assisted telephone interview; CCC = concordance correlation coefficient, CI = confidence interval; FE = fixed-effect; GH = general health perceptions, ICC = intra class correlation; IVR = interactive voice response; MD = mean difference; OR = odds ratio; PDA = Personal Digital Assistant; PF = physical functioning; PnP = paper and pencil; PRO = patient-reported outcome; RCT = randomized controlled trial; RE = random effect; RE-scale = role-emotional; RP = role-physical; SAQ = self-administered questionnaire, SMD = standardized mean difference

kursiv: eigene Berechnung/Interpretation

Anhang 8: Ergebnisse der systematischen Literaturrecherche: Antwortverhalten – Vergleiche elektronischer Erhebungsmodi

Tabelle A 12: Ergebnisse zur Übereinstimmung des Antwortverhaltens – Vergleiche elektronischer Erhebungsmodi

Erstautor/in (Jahr)	Eingeschlossene Studien	Analysierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
1 Smartphone vs web/PC							
Greenleaf et al. (2017)	1 crossover, 1 unclear	2.781	mobile phone web survey	web survey on PC	rate of reporting of sensitive behavior	no difference in report of sensitive behavior indicators (1 study); 2 of 5 questions were statistically significantly different (alcohol consumption, income higher in web/PC; 1 study)	-
Marcano Belisario et al. (2015)	1 crossover	90	app (iOS)	laptop	data equivalence (mean score differences in validated survey questionnaires)	no statistically significant difference	-

Erstautor/in (Jahr)	Eingeschlossene Studien	Analysierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
	1 crossover	90	app (iOS)	laptop	data equivalence (ICC)	exceeded treshold of 0.70	-

2 Tablet vs web/PC

Muehlhausen et al. (2015)	1 within subject (non-rand.)	122	tablet (touch)	web/PC	correlation	RE: 0.90 95 % CI [0.85, 0.95]	-
---------------------------	------------------------------	-----	----------------	--------	-------------	-------------------------------	---

3 IVRS vs web/PC

Muehlhausen et al. (2015)	1 rand. 3/4 group crossover	n.s.	web/PC	IVRS	correlation	RE: 0.88 95 % CI [0.84, 0.91]	-
---------------------------	-----------------------------	------	--------	------	-------------	-------------------------------	---

4 Smartphone vs PDA

Marcano Belisario et al. (2015)	1 paired repeated measure design	36	app (iOS)	PDA	data equivalence (correlation coefficient)	exceeded treshold of 0.60	0.60 interpreted as 'equivalent' by review author
---------------------------------	----------------------------------	----	-----------	-----	--	---------------------------	---

5 Smartphone vs SMS

Marcano Belisario et al. (2015)	1 crossover	48	app	SMS	data equivalence (mean score differences in non-validated survey questionnaires)	no statistically significant difference	-
	1 crossover	48	app	SMS	data completeness (mean number of daily entries)	statistically significantly differences for day 1, 2, 4, 5 and 6; no statistically	-

Erstautor/in (Jahr)	Eingeschlossene Studien	Analysierte Fälle	Modus 1	Modus 2	Endpunkt	Ergebnis	Kommentar
						significant difference for day 3	
6 PDA vs sonstige elektronische Erfassung							
Jibb et al. (2020)	1 rand. crossover	106	PDA	manually manipulated slide device-based tool	pain score equivalence (pearson correlation)	equivalence	-

Abkürzungen: CCC = concordance correlation coefficient; CI = confidence interval; ICC = intra class correlation; IVR = interactive voice response; PDA = Personal Digital Assistant; PnP = paper and pencil; PRO = patient-reported outcome; RCT = randomized controlled trial; RE = random effect; SAQ = self-administered questionnaire; SMD = standardized mean difference

kursiv: eigene Berechnung/Interpretation

Anhang 9: Unit-Nonresponse-Analyse im Rahmen der empirischen Untersuchung

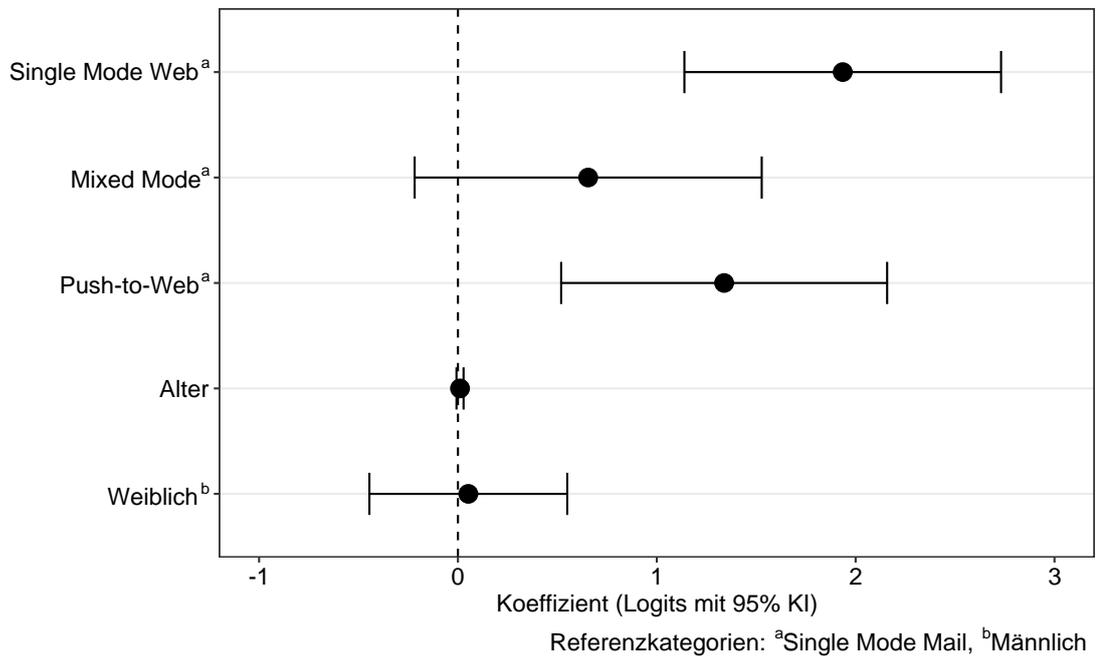


Abbildung A 1: Binomial-logistische Regression zur Unit-Nonresponse-Analyse

Anhang 10: Analyse der „Weiß nicht mehr“-Antworten im Rahmen der empirischen Untersuchung

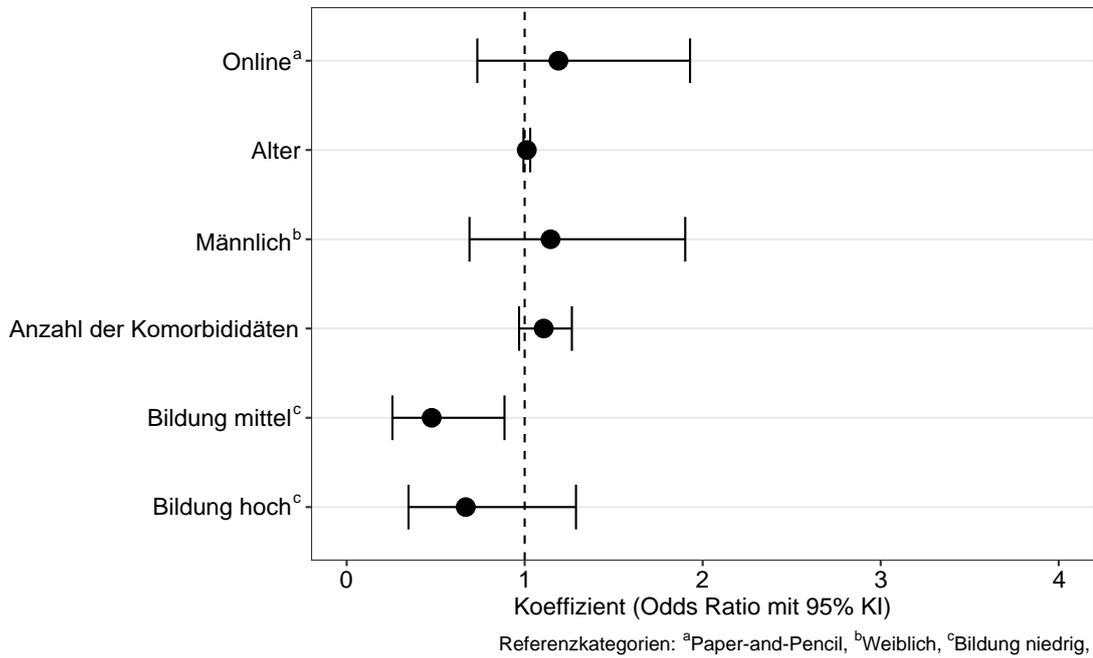


Abbildung A 2: Binomial-logistische Regression zur Analyse der „Weiß nicht mehr“-Antworten