

Abbildbarkeit des sozioökonomischen Status in der pharmakoepidemiologischen Forschungsdatenbank GePaRD: Beschreibung und Anwendung am Beispiel des Zusammenhangs mit Adipositas

Zusammenfassung

Hintergrund: Der sozioökonomische Status stellt bei vielen epidemiologischen Fragestellungen eine wichtige Einfluss- bzw. Störgröße dar. Dennoch wird diese Information in Studien basierend auf Versichertendaten oft nicht berücksichtigt, vermutlich auch da unklar ist, wie vollständig bzw. aussagekräftig die vorliegenden Informationen sind.

Ziel der Arbeit: Ziel war, den sozioökonomischen Status für die pharmakoepidemiologische Forschungsdatenbank GePaRD zu erfassen und zu bewerten.

Methodik: Zunächst wurde ein Algorithmus entwickelt, der den Versicherten im Jahr 2017 einen Schulbildungsstatus zuordnet und so den sozioökonomischen Status auf individueller Ebene schätzt. Zum Vergleich wurde der sozioökonomische Status basierend auf dem Wohnort geschätzt. Zur Plausibilisierung wurde untersucht, ob sich der bekannte Zusammenhang zwischen dem Vorliegen einer Adipositas und dem sozioökonomischen Status reproduzieren lässt.

Ergebnisse: Es konnte je nach Altersgruppe für 86–93% der unter 60-Jährigen und für 67% der 60- bis 69-Jährigen der sozioökonomische Status dichotom geschätzt werden. Für Personen ab 70 Jahren war der Anteil fehlender Werte sehr hoch. In allen Subgruppen war die Prävalenz der Adipositas bei Personen mit niedrigem sozioökonomischen Status höher als bei Personen mit höherem sozioökonomischen Status. Die wohnortbasierte Schätzung zeigte zwar auch plausible Ergebnisse, doch mit schwächeren Unterschieden nach sozioökonomischem Status.

Diskussion und Fazit: Insgesamt legen die Studienergebnisse nahe, dass sich die in GePaRD verfügbaren Informationen für eine Abschätzung des individuellen sozioökonomischen Status bei Altersgruppen bis 69 Jahren, was Plausibilität und Vollständigkeit betrifft, gut eignen.

Schlüsselwörter: sozioökonomischer Status, Bildung, sozialer Deprivationsindex, Versichertendaten, Adipositas

Marieke Asendorf¹

Jonas Reinold¹

Tania Schink¹

Bianca Kollhorst²

Ulrike Haug^{1,3}

1 Abteilung Klinische Epidemiologie, Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS, Bremen, Deutschland

2 Abteilung Biometrie und EDV, Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS, Bremen, Deutschland

3 Fachbereich Human- und Gesundheitswissenschaften, Universität Bremen, Deutschland

Einleitung

Wissenschaftliche Untersuchungen belegen seit Jahrzehnten, dass die Bürde verschiedenster Krankheiten sozial ungleich verteilt ist. In den meisten Fällen tragen dabei sozial benachteiligte Gruppen die größere Krankheitslast [1], [2], [3]. Ein ausgeprägter Zusammenhang besteht beispielsweise zwischen dem sozioökonomischen Status und dem Vorliegen einer Adipositas [4], [5], [6], [7], [8]. So berichten zum Beispiel Großschädl/Stronegger basierend auf Survey-Daten für Frauen mit grundlegendem Schulabschluss eine Adipositas-Prävalenz, die um etwa 13 Prozentpunkte höher ist als bei Frauen mit Hochschulreife. Für Männer wurde in dieser Studie ein Unterschied von etwa 9 Prozentpunkten beobachtet [9].

Dementsprechend wichtig ist es, Informationen zum sozioökonomischen Status in epidemiologischen Studien als mögliche Einfluss- bzw. Störgröße zu berücksichtigen. In Studien, die auf Primärdaten basieren, werden hierfür in der Regel Informationen zu schulischen und beruflichen Abschlüssen, zum Beruf oder zum Einkommen erhoben [6], [10]. Neben Primärdaten stellen Abrechnungsdaten von Krankenversicherungen eine zunehmend wichtige Datenquelle für die Versorgungsforschung und die epidemiologische Forschung dar. Die pharmakoepidemiologische Forschungsdatenbank (German Pharmacoepidemiological Research Database, kurz: GePaRD) beispielsweise enthält Abrechnungsdaten von vier gesetzlichen Krankenversicherungen, deckt etwa 20% der deutschen Bevölkerung ab und umfasst aktuell die Datenjahre 2004–2017.

GePaRD wird für zahlreiche Studien im Bereich der Arzneimittelanwendungs- und -risikoforschung und der Krebsfrüherkennung genutzt. Auch bei vielen dieser Studien ist die Berücksichtigung des sozioökonomischen Status von großer Relevanz bzw. Interesse, doch wurde bisher nicht systematisch untersucht, inwiefern in GePaRD vorhandene Variablen geeignet sind, um den sozioökonomischen Status abzubilden bzw. in welcher Vollständigkeit diese vorliegen.

Grundsätzlich gibt es in den Daten der gesetzlichen Krankenversicherungen verschiedene Möglichkeiten, den sozioökonomischen Status abzuschätzen (siehe Tabelle 1). In GePaRD bieten sich zwei Variablen an, die Informationen zum sozioökonomischen Status liefern könnten. Zum einen liegt in den Daten der Kreisschlüssel des Wohnorts der Versicherten vor, der – verknüpft mit dem regionalen Deprivationsindex des jeweiligen Landkreises – genutzt werden kann, um den sozioökonomischen Status anhand aggregierter Informationen abzuschätzen, d.h. allen Personen eines Landkreises wird der gleiche sozioökonomische Status zugeordnet. Zum anderen liegt eine Variable vor, die Informationen zum Beruf, zur beruflichen Stellung und zur Ausbildung der Versicherten gemäß Schlüsselverzeichnis der Bundesagentur für Arbeit enthält und damit eine Möglichkeit darstellt, den sozioökonomischen Status auf individueller Ebene abzuschätzen. Ziel dieser Studie war es, die Vollständigkeit dieser beiden Informationen in GePaRD zu untersuchen sowie deren Aussagekraft und Plausibilität bzgl. des sozioökonomischen Status indirekt zu bewerten, indem jeweils der Zusammenhang mit dem Vorliegen einer Adipositas ermittelt wurde.

Methoden

Datenquelle und Studienpopulation

Basis dieser Untersuchung war die pharmakoepidemiologische Forschungsdatenbank GePaRD. Sie enthält Abrechnungsdaten von vier gesetzlichen Krankenversicherungen in Deutschland und umfasst Informationen von derzeit ca. 25 Millionen Personen, die seit 2004 oder danach bei einer der teilnehmenden Krankenversicherungen versichert waren. Neben demographischen Angaben enthält GePaRD Informationen zu Arzneimittelverordnungen sowie zu ambulanten und stationären Leistungen und Diagnosen. Pro Datenjahr stehen Informationen zu ungefähr 20% der Allgemeinbevölkerung zur Verfügung und es sind alle geographischen Regionen Deutschlands vertreten. Für diese Auswertungen wurden die Daten der Jahre 2004 bis 2017 genutzt. Eingeschlossen wurden alle Versicherten aus GePaRD, die im Jahr 2017 mindestens an einem Tag versichert und an diesem Tag am Leben waren, für die eine Angabe zum Geschlecht, zum Geburtsjahr und zum Wohnort vorlag und deren Wohnort innerhalb Deutschlands lag.

Abschätzung des sozioökonomischen Status auf individueller Ebene

Ausgangspunkt war eine Variable, die in den Versichertenstammdaten enthalten ist und Informationen zum Beruf, zur beruflichen Stellung und zur Ausbildung enthält. Diese Informationen, die mit dem Begriff „Tätigkeitsschlüssel“ zusammengefasst werden, werden von den Arbeitgebenden gemäß §28a Sozialgesetzbuch (SGB) IV an die Sozialversicherungsträger für alle versicherungspflichtigen und freiwilligen Hauptversicherten mindestens jährlich übermittelt. Dementsprechend steht die Information unter anderem nicht für Personen zur Verfügung, die in dem jeweiligen Jahr bereits in Rente sind, sowie für sonstige Versicherte und Familienversicherte [11].

Die Operationalisierung dieser Variable wurde von folgenden Prinzipien geleitet:

A) Der Fokus sollte auf der Abschätzung des sozioökonomischen Status basierend auf den Informationen zur schulischen Bildung liegen.

B) Sofern eindeutig, wurden die Informationen zum beruflichen Abschluss oder zur beruflichen Stellung einbezogen, um Rückschlüsse auf die schulische Bildung zu ziehen. So wurde beispielsweise bei Personen mit universitärem Abschluss unabhängig von den Informationen zu ihrer schulischen Bildung von (fachgebundener) Hochschulreife ausgegangen.

C) Mit dem Ziel, den Personen, die im Jahr 2017 versichert waren, einen Bildungsstand zuzuordnen, wurde auch auf frühere Datenjahre zurückgegriffen. Bei verrenteten Personen beispielsweise konnte damit teilweise auf die Zeit der Berufstätigkeit zurückgegriffen werden, d.h. auf Jahre, in denen noch ein Tätigkeitsschlüssel übermittelt wurde.

D) Es wurde angenommen, dass Familienversicherte einen ähnlichen sozioökonomischen Status haben wie die zugehörigen Hauptversicherten. Auf Basis dieser Annahme wurde versucht, Personen, zu denen in keinem der Datenjahre ein Tätigkeitsschlüssel vorlag, mit einer familienangehörigen Person, der bereits ein Status zugeordnet werden konnte, über eine gemeinsame Versicherung in GePaRD zu verknüpfen. Für diese Verknüpfung wurde die ebenfalls in den Versichertenstammdaten enthaltene Familien-ID verwendet.

E) Lieferten die Daten unterschiedlicher Kalenderjahre diskrepante Informationen zur Bildung, wurde jeweils der höchste Bildungsstand herangezogen. Dem liegt die Annahme zugrunde, dass sich der erworbene Bildungsstand nicht verringert, sondern sich durch eine weitere schulische Ausbildung allenfalls erhöhen kann.

Basierend auf dieser Operationalisierung wurden die Versicherten in die folgenden drei Kategorien eingeteilt:

1. (Fachgebundene) Hochschulreife,
2. Hauptschulabschluss oder mittlere Reife,
3. Bildung unbekannt.

Tabelle 1: Übersicht über Informationen zur Abschätzung des sozioökonomischen Status in Daten der gesetzlichen Krankenkassen

Dimension	ggf. Einschränkungen	Verfügbarkeit in GePaRD
Beitragsgruppe		
definiert die Zugehörigkeit zu unterschiedlichen Beitragsgruppen [25]		nein
Versichertengruppe als Kombination aus Versichertenart und Personengruppenschlüssel		
ermöglicht u.a. eine Unterscheidung von freiwillig versicherten Arbeitnehmenden, Personen mit Niedriglohtätigkeit mit Sozialleistungen und Renten-Beziehenden bzw. Pensionärinnen und Pensionären [26]		nein
ausgeübte Tätigkeit		
definiert die ausgeübte Tätigkeit nach der Klassifikation der Berufe der Bundesagentur für Arbeit	nur während einer Berufstätigkeit verfügbar (Hauptversicherte)	ja
allgemeinbildender Schulabschluss		
definiert den höchsten allgemeinbildenden Schulabschluss	nur während einer Berufstätigkeit verfügbar (Hauptversicherte)	ja
beruflicher Ausbildungsabschluss		
definiert den höchsten beruflichen Ausbildungsabschluss	nur während einer Berufstätigkeit verfügbar (Hauptversicherte)	ja
Entgelt		
gibt das Einkommen für die Beitragsberechnung an [25]	nur bis zur Beitragsbemessungsgrenze verfügbar; nur für sozialversicherungspflichtig Beschäftigte verfügbar	nein
Wohnort mit Postleitzahl		
kann in Kombination mit regionalen Indikatoren genutzt werden [25]		auf Kreisebene verfügbar

In der letzten Kategorie wurden die Versicherten mit den Ausprägungen „Bildung unbekannt“ und „kein Abschluss“ sowie mit fehlendem Wert bzgl. schulischer Bildung zusammengefasst, da erwartet wurde, dass die Gruppe „kein Abschluss“ klein ist. Diese Einteilung orientiert sich an den Möglichkeiten, die der Tätigkeitsschlüssel bot, bevor er 2011 auf ein neues Klassifizierungssystem umgestellt wurde [11], [12], [13]. Das neue Klassifizierungssystem bietet zusätzlich die Möglichkeit, zwischen Hauptschulabschluss und mittlerer Reife zu unterscheiden (Abbildung 1). An GePaRD wurde der Tätigkeitsschlüssel gemäß der neuen Klassifikation jedoch größtenteils erst ab dem Datenstand von 2017 übermittelt. Um auch das Potenzial des neuen Tätigkeitsschlüssels zu untersuchen, führten wir Sensitivitätsanalysen durch. Hierfür wurde die Population der Hauptanalyse auf Personen beschränkt, deren Bildungsstatus auf Informationen über die berufliche oder schulische Bildung des Klassifikationssystems nach 2011 basiert.

Abschätzung des sozioökonomischen Status anhand der Informationen zum Wohnort

Die aktuellsten in GePaRD zur Verfügung stehenden Informationen zum Wohnort auf Kreisebene wurde mit dem

vom Robert Koch-Institut (RKI) entwickelten German Index of Socioeconomic Deprivation (GISD) verknüpft, wobei die Revision dieses Index verwendet wurde, die auf Daten von 2014 basiert. Der Index gibt das Ausmaß der Deprivation basierend auf den drei Dimensionen Einkommen, Bildung und Beruf wieder [14]. Passend zur kleinsten räumlichen Einheit in GePaRD wurde der GISD auf Kreisebene verwendet (weitere Erläuterungen im Appendix (Anhang 1). Analog zum Vorgehen von Kroll et al. wurden die den Kreisen zugeordneten Deprivations-Scores in Quintile eingeteilt (erstes Quintil: „niedrige Deprivation“; zweites bis viertes Quintil: „mittlere Deprivation“; fünftes Quintil: „hohe Deprivation“) [14].

Erfassung von Informationen zum Vorliegen einer Adipositas

Die Erfassung des Vorliegens einer Adipositas basierte zum einen auf den hierfür relevanten ICD-10-Codes (stationär oder gesichert ambulant) sowie auf Codes für Prozeduren, die auf eine Therapie einer Adipositas hinweisen (z.B. magenverkleinernde chirurgische Eingriffe). Eine Person wurde als adipös klassifiziert, wenn für sie in GePaRD mindestens einer der relevanten Codes vorlag. Dabei wurde nicht nur das Datenjahr 2017 betrachtet, sondern auch alle in GePaRD verfügbaren Datenjahre zu

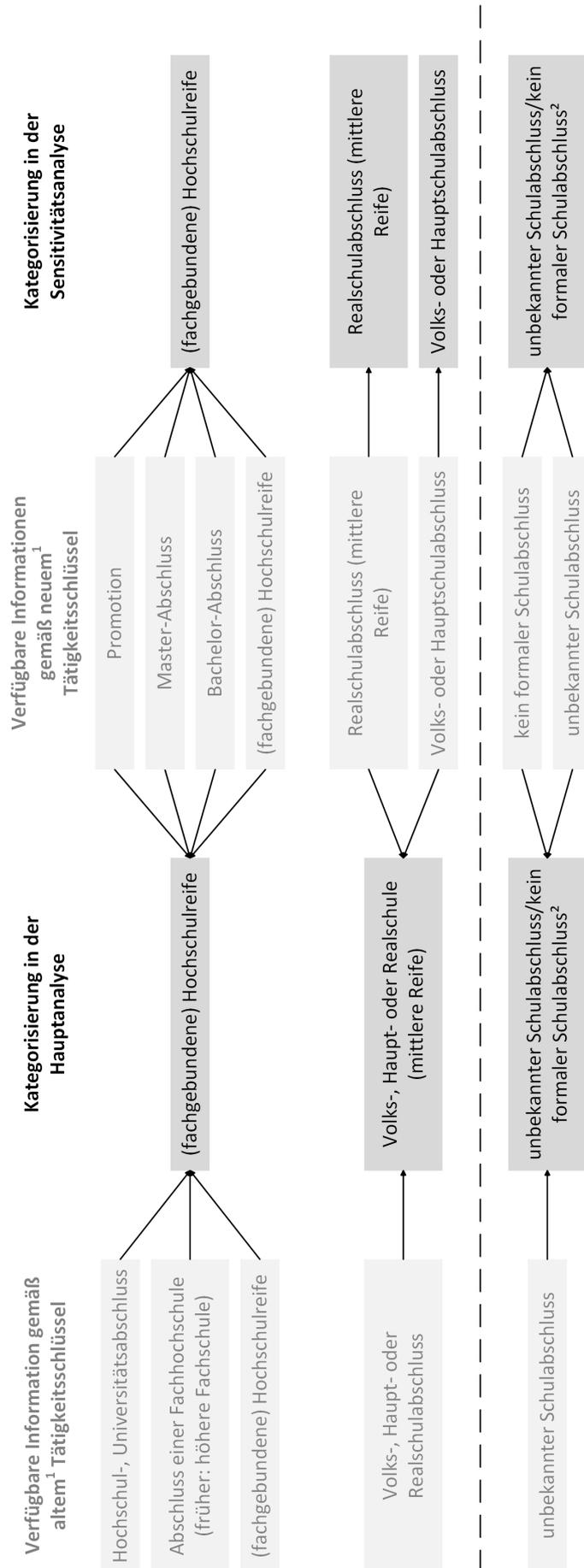


Abbildung 1: Kategorisierung der Informationen zur Bildung in GePaRD

¹ Die Bezeichnung alter Tätigkeitsschlüssel bezeichnet den Tätigkeitsschlüssel vor der Umstellung 2011 durch die Bundesagentur für Arbeit. Die Bezeichnung neuer Tätigkeitsschlüssel bezeichnet den Tätigkeitsschlüssel nach der Umstellung. In GePaRD sind erstmals mit der Datenlieferung für 2017 für alle Krankenkassen Daten zum neuen Tätigkeitsschlüssel (auch retrospektiv) geliefert worden.

² Die Kategorie „unbekannter Schulabschluss / kein formaler Schulabschluss“ wurde in den Analysen zum Zusammenhang mit der Adipositas-Prävalenz nicht berücksichtigt.

dieser Person vor 2017. Dem liegt die Überlegung zugrunde, dass es sich bei Adipositas in der Regel um eine chronische und damit dauerhaft vorliegende Gesundheitsstörung handelt. Da sie vermutlich oft nur codiert wird, wenn dies als Abrechnungsgrund relevant ist, was nicht jedes Jahr der Fall sein muss, erhöht sich durch das Zurückgreifen auf frühere Datenjahre die Sensitivität der Erfassung der Adipositas.

Auswertung der Daten in GePaRD

Zunächst wurde zur Beurteilung der Vollständigkeit der individuellen Information zur schulischen Bildung in GePaRD der Anteil der Versicherten im Jahr 2017 ermittelt, denen basierend auf dem oben beschriebenen Vorgehen eine schulische Bildung zugeordnet werden konnte. Zur Ermittlung der Adipositas-Prävalenz wurden die Versicherten des Jahres 2017, bei denen gemäß oben beschriebenen Vorgehen eine Adipositas vorlag, im Zähler betrachtet und alle Personen aus der beschriebenen Versichertenkohorte im Nenner. Die so berechnete Adipositas-Prävalenz wurde in den Hauptanalysen gemäß der dichotomen Klassifikation stratifiziert nach (Fach-)Hochschulreife vs. Real- oder Hauptschulabschluss ermittelt. In der Sensitivitätsanalyse wurde die Prävalenz stratifiziert nach den drei Kategorien (Fach-)Hochschulreife, mittlere Reife und Hauptschulabschluss ermittelt. Für die Auswertungen basierend auf dem GISD wurde analog vorgegangen, d.h. die Adipositas-Prävalenz wurde stratifiziert nach den Kategorien niedrige, mittlere und hohe sozioökonomische Deprivation ermittelt. Alle Analysen wurden hinsichtlich Alter und Geschlecht der Versicherten stratifiziert.

Vergleichende Auswertung von Primärdaten

Ergänzend zu den Auswertungen basierend auf GePaRD wurden die im Rahmen der GEDA-2014/2015-Befragung des RKI erhobenen Primärdaten vergleichend ausgewertet [15], d.h. es wurde auch die Adipositas-Prävalenz nach Alter, Geschlecht und Schulbildung (in den Kategorien (Fach-)Abitur, mittlere Reife und Hauptschulabschluss) ermittelt. An der Befragung, die mittels elektronischer Fragebögen oder Telefoninterviews durchgeführt wurde, nahmen 24.016 Personen ab 18 Jahren mit Hauptwohnsitz in Deutschland teil [16].

Ergebnisse

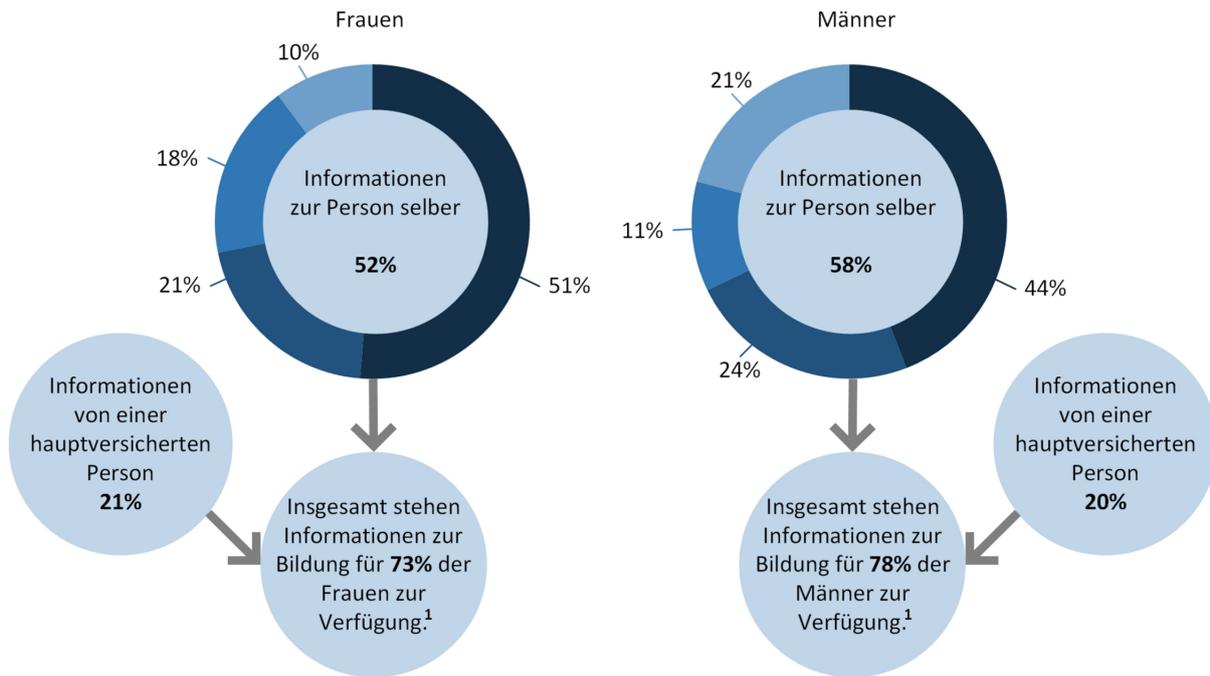
Nach Anwendung der Ein- und Ausschlusskriterien (258.013 Personen wurden aufgrund fehlender Informationen ausgeschlossen; davon hatten 217.600 Personen keine gültige Angabe zum Wohnort in Deutschland) wurden 17.317.559 Personen in die Untersuchung eingeschlossen (Anteil Frauen: 53%). Insgesamt konnte für 73% der Frauen und 78% der Männer der sozioökonomi-

sche Status auf individueller Ebene gemäß dem beschriebenen Algorithmus approximiert werden.

Bei etwa der Hälfte der Männer und Frauen lagen Informationen zum eigenen Bildungsabschluss vor. Bei weiteren ca. 20% lag zwar keine eigene Bildungsinformation vor, aber es konnten Personen aus der Familie mit vorhandener Bildungsinformation zugeordnet werden und damit konnte der sozioökonomische Status indirekt abgeschätzt werden (siehe Abbildung 2). Der Anteil der Versicherten, denen über den Algorithmus ein sozioökonomischer Status zugeordnet werden konnte, lag in den Altersgruppen unter 60 Jahren bei Männern und Frauen durchweg über 85%, bei 60–69-Jährigen bei 67% (Frauen: 63%, Männer: 72%), bei 70–79-Jährigen bei 16% (Frauen: 10%, Männern: 23%) und bei Personen ab 80 Jahren unter 1% (Frauen: 0,3%, Männer: 1%) (siehe Tabelle 2). Der Anteil ohne Bildungsabschluss an allen Versicherten betrug über alle Subgruppen hinweg weniger als 1%.

Abbildung 3 zeigt die in GePaRD beobachtete Adipositas-Prävalenz stratifiziert nach Altersgruppen, Geschlecht und zugeordnetem Bildungsstand (Hauptschulabschluss/mittlere Reife und (Fach-)Hochschulreife) auf individueller Ebene. In allen Altersgruppen und bei beiden Geschlechtern war die Prävalenz der Adipositas, die insgesamt mit dem Alter anstieg, bei Personen mit niedrigerem Bildungsstand höher. Die Punktschätzer der Prävalenz bei Frauen mit 30-39 Jahren waren beispielsweise ca. 9 Prozentpunkte über denen der Personen mit höherer Bildung (22% vs. 13%). In der gleichen Altersgruppe betrug die Differenz bei den Männern ca. 5 Prozentpunkte (12% vs. 7%). In Abb. 7 (Anhang 1) ist die Adipositas-Prävalenz für alle Kategorien inkl. der Kategorien „Bildung unbekannt“, kein formaler Bildungsabschluss und Personen ohne jegliche Angabe zum Bildungsabschluss vergleichend dargestellt. In Abb. 8 (Anhang 1) wird die Adipositas-Prävalenz außerdem für die zwei Gruppen mit bzw. ohne Information zur Bildung verglichen. Dabei zeigt sich bei Männern in beiden Gruppen eine weitgehend übereinstimmende Prävalenz, außer bei den 18- bis 29-Jährigen (niedrigere Adipositas-Prävalenz in der Gruppe mit fehlender Information zur Bildung). Bei Frauen zeigt sich bis zum Alter von 50 Jahren ein ähnliches Muster wie bei Männern. Bei älteren Frauen zeigt sich in der Gruppe mit Information zur Bildung eine höhere Adipositas-Prävalenz als bei Frauen ohne Information zur Bildung.

In die Sensitivitätsanalysen zur Einteilung des Bildungsstands in drei Kategorien (Hauptschulabschluss, Real-schulabschluss und (Fach-)Abitur) konnten 27% der Versicherten eingeschlossen werden (24% der Männer, 29% der Frauen). Es zeigte sich in allen Subgruppen unter 80 Jahren ein inverser Zusammenhang zwischen Adipositas-Prävalenz und Bildungsstatus, wobei der Unterschied zwischen den Kategorien Realschul- vs. Hauptschulabschluss kleiner war (beispielsweise 5% Prozentpunkte bei den 30- bis 39-jährigen Frauen) als der Unterschied zwischen den Kategorien Realschulabschluss vs. (Fach-)Hochschulreife (8 Prozentpunkte bei den 30- bis 39-jährigen Frauen) (Abbildung 4). Der maximale Unterschied in der Adipositas-Prävalenz zwischen Personen



Aufteilung der spezifischen Datenquellen zur individuellen Person (als relative Werte sind die Anteile der Personen eines Geschlechts angegeben, deren Bildungsinformation aus der entsprechenden Quelle stammt):

- schulische Bildung (basierend auf dem neuen Beschäftigungsschlüssel)
- akademischer beruflicher Abschluss
- Beschäftigungsniveau
- schulische Bildung (basierend auf dem alten Beschäftigungsschlüssel)

¹ Dabei wird der Code „kein Schulabschluss“ als gültiger Code klassifiziert und ist entsprechend in dieser Zahl enthalten.

Abbildung 2: Quelle der Bildungsinformationen

Tabelle 2: Anteil der Versicherten, denen ein sozioökonomischer Status zugeordnet werden konnte, nach Alter und Geschlecht (dabei wird der Code „kein Schulabschluss“ als gültiger Code klassifiziert und ist entsprechend in den berechneten Anteilen enthalten)

Alter	Zuordnung eines sozioökonomischen Status		
	Beide Geschlechter	Frauen	Männer
Jünger als 18 Jahre	88,4%	88,5%	88,4%
18–29 Jahre	88,2%	88,7%	87,7%
30–39 Jahre	92,6%	93,3%	92,0%
40–49 Jahre	89,5%	90,3%	88,6%
50–59 Jahre	85,7%	86,1%	85,3%
60–69 Jahre	67,0%	63,4%	71,6%
70–79 Jahre	15,7%	10,0%	23,3%
80 Jahre und älter	0,7%	0,3%	1,2%

mit Hauptschulabschluss vs. Personen mit (Fach-)Hochschulreife lag bei 13 Prozentpunkten und wurde in der Subgruppe der 30- bis 39-jährigen Frauen beobachtet. Die Ergebnisse zum Zusammenhang zwischen der in GePaRD beobachteten Adipositas-Prävalenz und dem mittels räumlichem Deprivationsindex des Wohnorts (Kreisebene) abgeschätzten sozioökonomischen Status

sind in Abbildung 5 dargestellt. Auch hier zeigt sich in allen Altersgruppen und für beide Geschlechter eine Zunahme der Adipositas-Prävalenz mit steigender Deprivation. Die Unterschiede sind im Vergleich zur Auswertung basierend auf dem individuell abgeschätzten sozioökonomischen Status weniger ausgeprägt. Der Unterschied zwischen den Subgruppen mit der höchsten bzw. der nied-

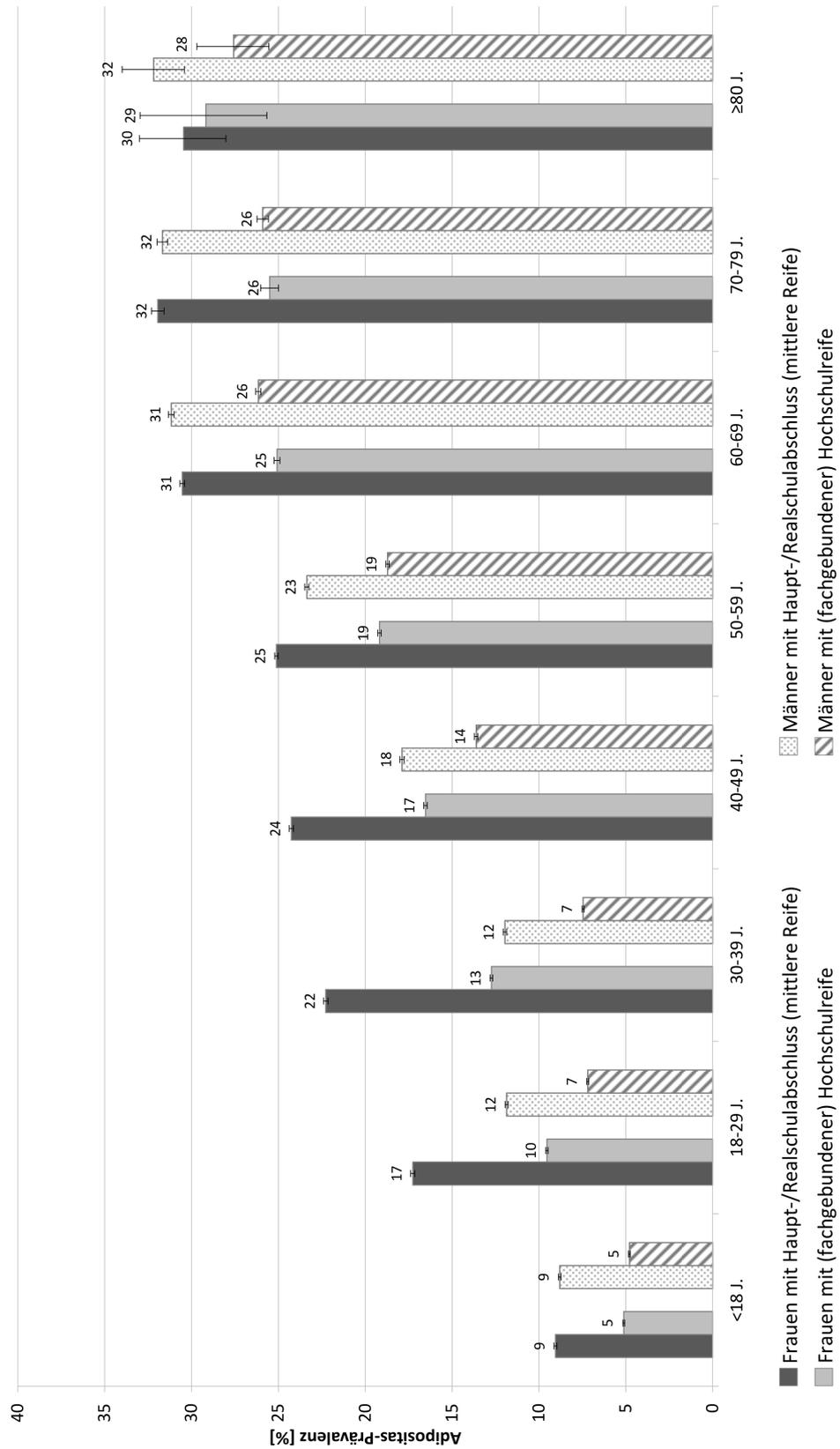


Abbildung 3: Prävalenz von Adipositas in GePaRD für 2017 stratifiziert nach Alter, Geschlecht und individuell zugeordnetem Bildungsstand

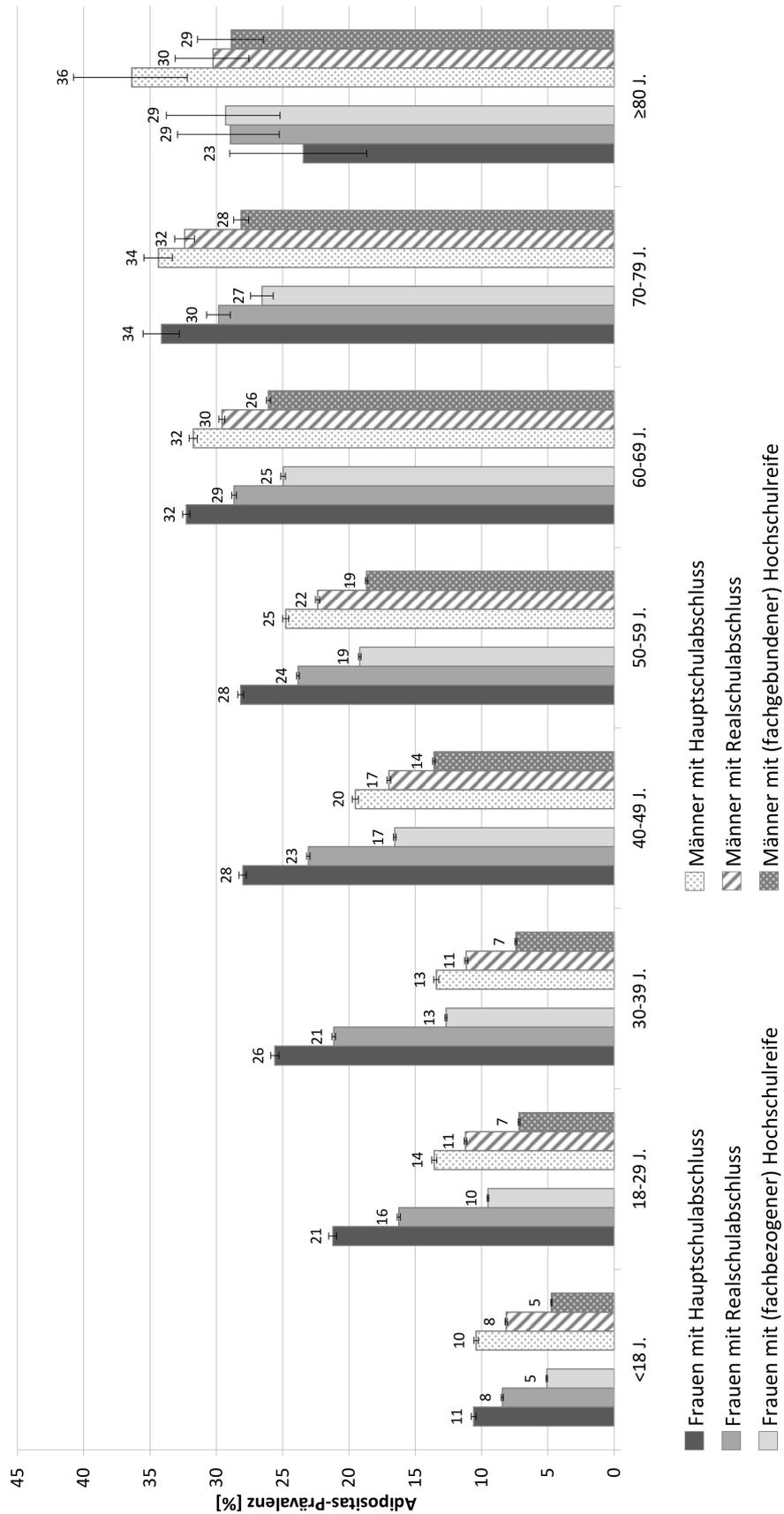


Abbildung 4: Prävalenz von Adipositas in GePaRD für 2017 stratifiziert nach Alter, Geschlecht und individuell zugeordnetem Bildungsstand (in drei Kategorien basierend auf neuem Berufsschlüssel)

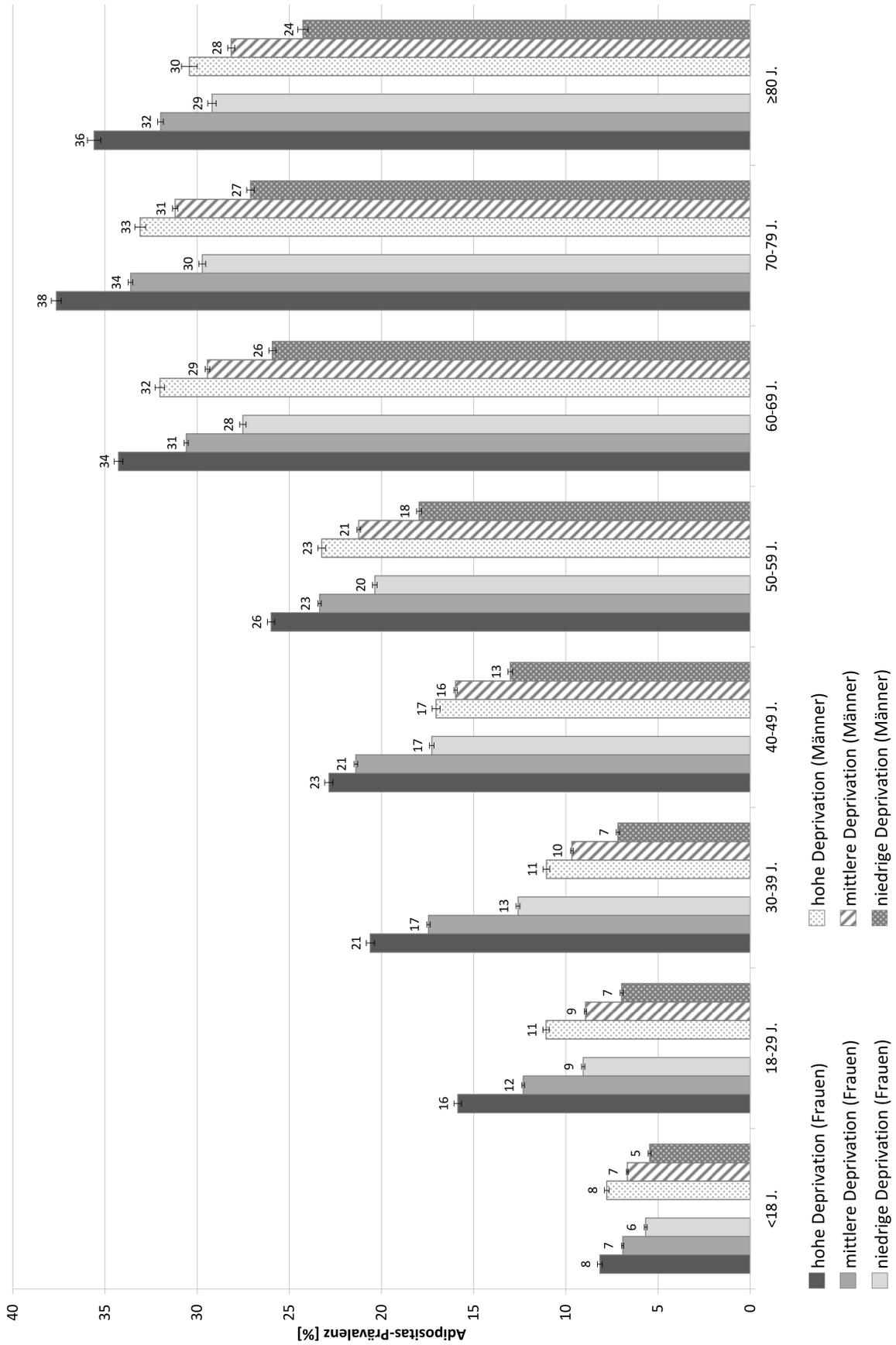


Abbildung 5: Prävalenz von Adipositas in GePaRD für 2017 stratifiziert nach Alter, Geschlecht und regionalem Deprivationsgrad (GISD)

rigsten Deprivation beträgt bei den 30- bis 39-jährigen Frauen ca. 8 Prozentpunkte (21% vs. 13%) und bei den Männern dieser Altersklasse ca. 4 Prozentpunkte (11% vs. 7%).

Die Ergebnisse der Vergleichsanalyse auf Basis des GEDA-2014/2015-Datensatzes sind in Abbildung 6 dargestellt. Dieser Datensatz enthält Informationen zum Bildungsstand auf individueller Ebene, d.h. die Ergebnisse können den in Abbildung 3 dargestellten Ergebnissen zu GePaRD gegenübergestellt werden. Im Vergleich zu GePaRD zeigte sich in einigen Subgruppen – vor allem in höheren Altersgruppen – eine etwas andere, oft niedrigere Adipositas-Prävalenz, aber insgesamt waren die Muster, was die Unterschiede zwischen den Bildungsgruppen betrifft, ähnlich. Tendenziell war der absolute Unterschied in der Adipositas-Prävalenz zwischen den Bildungsgruppen bei GEDA stärker ausgeprägt und zeigte weniger Geschlechtsunterschiede als bei GePaRD. Bei den 40- bis 49-jährigen Frauen beispielsweise betrug der Unterschied zwischen der höheren und der niedrigeren Bildungsgruppe 7 Prozentpunkte (GePaRD: 7 Prozentpunkte) und bei 40- bis 49-jährigen Männern betrug der Unterschied 8 Prozentpunkte (GePaRD: 4 Prozentpunkte).

Diskussion

Die Ergebnisse dieser Studie zeigen, dass es anhand von individuellen Informationen, die in der pharmakoepidemiologischen Forschungsdatenbank GePaRD verfügbar sind, möglich ist, etwa drei Viertel der Versicherten in eine höhere bzw. niedrigere Bildungskategorie als Annäherung an den sozioökonomischen Status einzuteilen. Die Tatsache, dass basierend auf dieser Zuordnung der bekannte Zusammenhang hinsichtlich Adipositas und sozioökonomischem Status reproduziert werden konnte, bestätigt die Plausibilität dieser Klassifikation. Auch für die Abschätzung des sozioökonomischen Status anhand des Landkreises des Wohnorts zeigten sich insgesamt plausible Ergebnisse. Der in diesen Analysen beobachtete Unterschied zwischen den 20% Versicherten mit dem niedrigsten bzw. dem höchsten Deprivations-Score (d.h. 60% der Versicherten bleiben bei dem Vergleich unberücksichtigt) entsprach jedoch im Ausmaß dem Unterschied, der bei der individuellen Abschätzung des sozioökonomischen Status bereits bei einer nur dichotomen Einteilung aller Versicherten mit vorhandener Bildungsinformation beobachtet wurde. Bei Einteilung des Bildungsstands in drei Kategorien waren die Unterschiede dann noch deutlich ausgeprägter, was insgesamt auch dafür spricht, dass die Trennschärfe der Abschätzung des sozioökonomischen Status basierend auf individuellen Informationen größer ist als bei Abschätzung basierend auf dem Landkreis des Wohnorts.

Die Stärke des Zusammenhangs zwischen dem Vorliegen einer Adipositas und dem sozioökonomischen Status sowie die Unterschiede nach Alter und Geschlecht waren insgesamt konsistent mit den Ergebnissen der Vergleichsanalysen, die wir basierend auf den GEDA-Daten durch-

geführt haben. Für einige Subgruppen war der Unterschied nach sozioökonomischem Status in den GEDA-Analysen etwas stärker ausgeprägt als in den GePaRD-Analysen. Bei den 50- bis 59-Jährigen beispielsweise zeigten die GEDA-Analysen für Personen mit Haupt-/Realschulabschluss im Vergleich zur höheren Bildungsgruppe eine zweifach (Frauen) bzw. 1,5-fach (Männer) erhöhte Adipositas-Prävalenz, während die GePaRD-Analysen eine 1,4-fach höhere Prävalenz für Männer und Frauen zeigten. Beim Vergleich der Ergebnisse ist jedoch zu berücksichtigen, dass in den GEDA-Analysen die Schätzer zum Teil weite Konfidenzintervalle haben. Auch ist zu bedenken, dass schon wegen der unterschiedlichen Art der Erfassung der Adipositas keine volle Übereinstimmung zu erwarten war. In Versicherungsdaten erscheinen Diagnosen vor allem dann, wenn sie abrechnungstechnisch relevant sind, d.h. wenn Leistungen im Zusammenhang mit der Diagnose erbracht wurden. In den GePaRD-Analysen wurde eine sensitive Faldefinition angewendet, was zu einer Überschätzung der Prävalenz im Erwachsenenalter geführt haben könnte. Bei der Definition von Adipositas in GEDA, die auf Eigenangaben zu Körpergröße und Gewicht basiert, gibt es andere Gründe, die zu einer Missklassifikation geführt haben könnten. So ist bekannt, dass Körpergewicht und -größe in Befragungen oft falsch angegeben werden [17]. Laut Maukonen et al. führt diese Missklassifikation in europäischen Studien zu einer Unterschätzung der Adipositas-Prävalenz um 0,6 bis 8,4 Prozentpunkte [17]. Somit enthält keine der beiden Datenquellen Goldstandard-Informationen hinsichtlich des Vorliegens einer Adipositas, weshalb Unterschiede in der Stärke des Zusammenhangs nicht überinterpretiert werden sollten. Dies gilt auch für den Vergleich mit anderen Studien, die aber insgesamt ebenfalls ähnliche Zusammenhänge zwischen dem individuell abgeschätzten sozioökonomischen Status und dem Vorliegen einer Adipositas zeigen [6], [7], [9].

Die Abschätzung des individuellen sozioökonomischen Status in unseren Analysen basierte auf den in GePaRD verfügbaren Informationen zum Tätigkeitsschlüssel. Grundsätzlich wären in Versichertendaten auch Informationen vorhanden, die eine noch genauere Abschätzung des sozioökonomischen Status ermöglichen würden, wie etwa die Beitragshöhe, die auf das Einkommen schließen lässt, oder der genaue Wohnort. Da es sich hierbei jedoch um sehr sensible Informationen handelt, stehen sie für kassenexterne Sekundärdatennutzung in der Regel nicht zur Verfügung. Die meisten der bisherigen Studien basierend auf Kassendaten, die den sozioökonomischen Status berücksichtigten, nutzten deshalb (grobe) Informationen zum Wohnort. Eine ältere Studie von Geyer basierend auf Daten von 416.000 Versicherten der AOK Mettmann aus den Jahren 1987–1996 zeigte zwar bereits das Potenzial der Nutzung von Informationen zu Bildung und Beruf zur Abschätzung des individuellen sozioökonomischen Status in Versichertendaten [1], doch fand dies in neueren Studien nur wenig Berücksichtigung. Dies könnte zum einen daran liegen, dass den Datenauswertenden der Tätigkeitsschlüssel nicht zur Verfügung stand, zum anderen ist

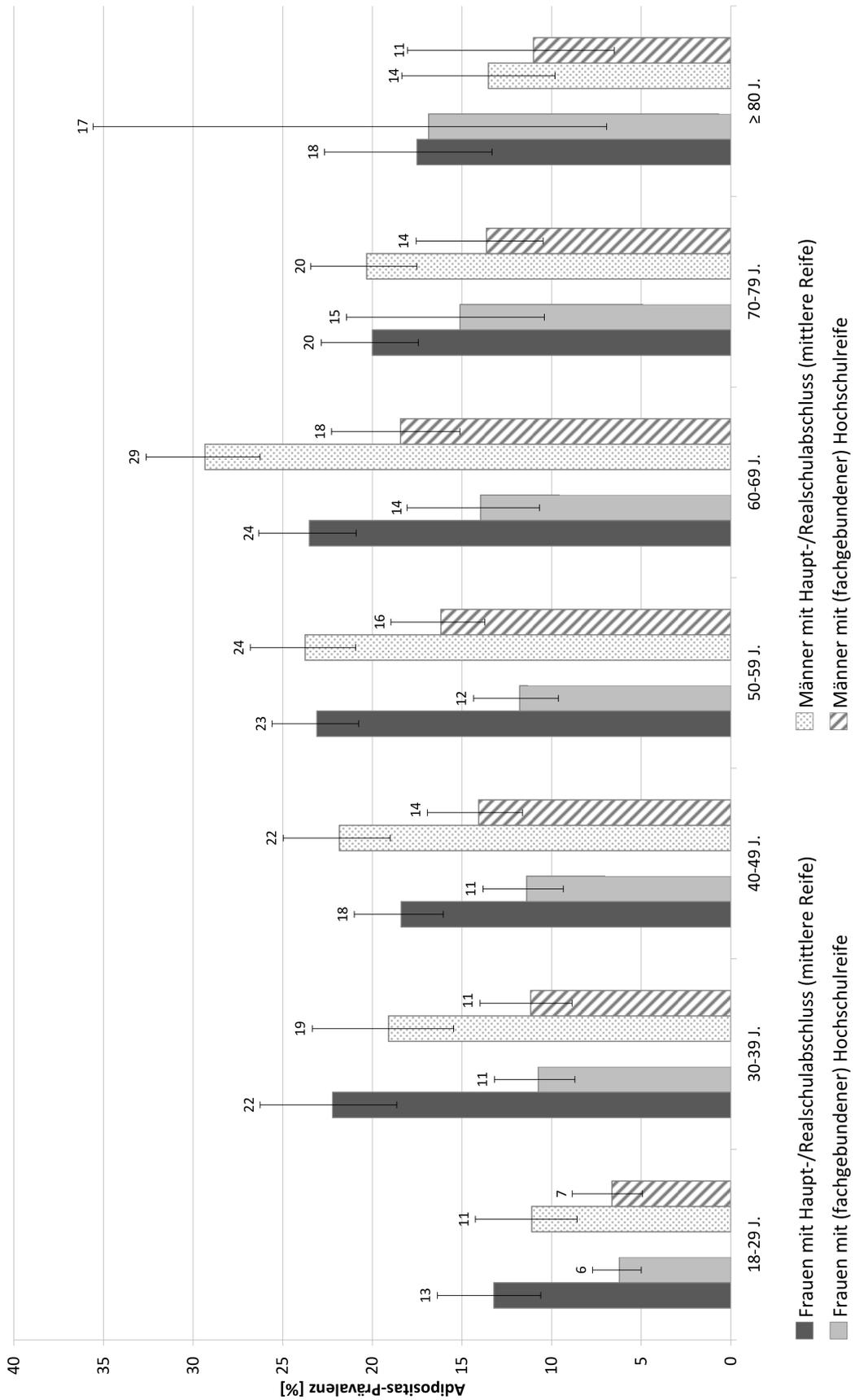


Abbildung 6: Prävalenz von Adipositas in GEDA für 2014–2015 stratifiziert nach Alter, Geschlecht und individuell zugeordnetem Bildungsstand

aber auch die Ansicht verbreitet, dass die Informationen nicht in der notwendigen Vollständigkeit bzw. Eindeutigkeit vorliegen [18]. In den GePaRD-Analysen konnten wir die Vollständigkeit durch das Zurückgreifen auf den langen Vorbeobachtungszeitraum in Verbindung mit der Verknüpfung von Familienversicherten sowie weiteren Rückschlüssen und Annahmen deutlich erhöhen. Eine dieser Annahmen, nämlich die Übertragbarkeit des sozioökonomischen Status von Hauptversicherten auf Familienangehörige, wurde zwar in der Literatur kontrovers diskutiert [19], [20], [21], doch ist sie aus unserer Sicht gut zu vertreten und basiert auf ähnlichen Annahmen wie die Verwendung des Familieneinkommens in primärdatenbasierten Studien.

Mit GePaRD als Datenquelle in Kombination mit der hier angewendeten Methodik ist es gelungen, für 86–93% der unter 60-Jährigen den individuellen sozioökonomischen Status grob (Hauptschulabschluss/mittlere Reife und (Fach-)Hochschulreife) abzuschätzen. Auch bei den 60- bis 69-Jährigen lag dieser Anteil noch bei 67% (Männer: 72% Frauen: 63%). Es ist außer Frage, dass es sich – aus mehreren Gründen – hierbei nicht um eine perfekte Klassifikation des sozioökonomischen Status handelt. Abgesehen von den zugrundeliegenden Annahmen kann auch nicht immer davon ausgegangen werden, dass die Meldung der Beschäftigungsschlüssel durch die Arbeitgebenden grundsätzlich fehlerfrei erfolgt [11]. Außerdem deckt der Bildungsstand nicht alle Dimensionen des sozioökonomischen Status ab [22]. Für zahlreiche deskriptive oder analytische Fragestellungen, die basierend auf Versichertendaten bearbeitet werden, ist jedoch auch diese grobe Abschätzung des sozioökonomischen Status von großem Wert. Dies umso mehr, als er auch als Ersatzparameter für sonstige Faktoren, die in Versichertendaten schlecht abgebildet sind (z.B. Lebensstilfaktoren), Bedeutung hat. So kann der derart geschätzte sozioökonomische Status zur Stratifizierung und Adjustierung genutzt werden, um die Relevanz als Einflussfaktor oder Störfaktor zumindest ansatzweise zu erkennen bzw. zu berücksichtigen. Bei der Einordnung der oben genannten Limitationen ist außerdem zu bedenken, dass auch die Schätzung des sozioökonomischen Status basierend auf Primärdaten u.a. aufgrund fehlender Angaben oft suboptimal ist.

Bei der Verwendung des GISD zur Schätzung des sozioökonomischen Status war der bekannte Zusammenhang zwischen niedrigem sozioökonomischen Status (bzw. hoher Deprivation) und Adipositas schwächer ausgeprägt als bei der individuellen Schätzung des sozioökonomischen Status in drei Kategorien, was nicht überraschend war [23], [24]. Da der Wohnort in GePaRD auf Kreisebene vorliegt, wird auch der GISD mit diesem Raumbezug verwendet. Damit können für Deutschland 401 regionale Einheiten unterschieden werden. Diese Einteilung ist sehr grob und trägt den Unterschieden im sozioökonomischen Status zwischen Individuen einer Region keine Rechnung, d.h. es ist von einer erheblichen Missklassifikation auszugehen. Dementsprechend ist die Schätzung des individuellen sozioökonomischen Status in GePaRD sowohl

dichotom als auch in drei Kategorien gegenüber dem GISD vorzuziehen. Der GISD könnte als zweitbeste Option in Betracht gezogen werden, wenn der individuelle sozioökonomische Status nicht geschätzt werden kann, was bei Personen ab 70 Jahren derzeit häufig der Fall ist, da bei ihnen oft nicht auf die Zeit der Berufstätigkeit zurückgegriffen werden kann. Es ist aber abzusehen, dass perspektivisch, d.h. mit noch längerer Beobachtungszeit in GePaRD, der Anteil mit fehlendem Tätigkeitsschlüssel auch bei den Älteren weiter abnimmt. Gleichmaßen ist abzusehen, dass für immer mehr Personen Bildungs- und Berufsinformationen nach dem neuen Klassifikationssystem vorliegen werden, d.h. in Zukunft wäre so, besonders für junge Menschen, eine Einteilung in drei Bildungskategorien möglich.

Der Vergleich der Adipositas-Prävalenz bei Personen mit vs. ohne Information zur Bildung (Abb. 8 (Anhang 1)) zeigt, dass insbesondere bei 18- bis 29-jährigen Männern und Frauen sowie bei Frauen ab 50 Jahren Unterschiede zwischen beiden Gruppen vorliegen. In diesen Gruppen scheint somit ein Zusammenhang zwischen dem Vorhandensein von Informationen zur Bildung und dem Sozialstatus zu bestehen. Bei den 18- bis 29-Jährigen erscheint dies plausibel, da in dieser Gruppe vermutlich Studierende enthalten sind (höherer Bildungsstand), die noch über ihre Eltern versichert sind. Da die Eltern zu dem Zeitpunkt teilweise bereits das Rentenalter erreicht haben und damit der Tätigkeitsschlüssel fehlt, ergeben sich auch für die Mitversicherten fehlende Werte. Personen, die hingegen eine Berufsausbildung durchlaufen (tendenziell niedrigerer Bildungsstand) beziehen in diesem Alter schon ein eigenes Gehalt und sind selbst versichert, d.h. es liegt ein Tätigkeitsschlüssel vor. Das Muster bei älteren Frauen lässt sich möglicherweise dadurch erklären, dass Frauen mit niedrigerem Bildungsstand häufiger keinen eigenen Beruf ausüben, sondern bei ihrem Ehemann mitversichert sind, der bereits in Rente ist (d.h. häufig fehlender Tätigkeitsschlüssel). Aufgrund dieser Zusammenhänge sollte in zukünftigen Auswertungen die Kategorie „missing“ gesondert betrachtet werden, v.a. was die 18- bis 29-Jährigen sowie Frauen über 50 Jahren betrifft.

Fazit

Insgesamt zeigen die hier dargestellten Studienergebnisse, dass sich die in GePaRD verfügbaren Informationen für eine dichotome Abschätzung des individuellen sozioökonomischen Status bei Altersgruppen bis 69 Jahren gut eignen, sowohl was die Vollständigkeit als auch was die Plausibilität betrifft. Bei Personen ab 70 Jahren könnte die Abschätzung des sozioökonomischen Status basierend auf dem GISD als zweitbeste Option genutzt werden, bis – mit noch längerer Beobachtungszeit in GePaRD – auch für diese Altersgruppen eine höhere Vollständigkeit erreicht sein wird.

Anmerkungen

Interessenkonflikte

Die Autoren und Autorinnen erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Marieke Asendorf, Jonas Reinold, Tania Schink, Bianca Kollhorst und Ulrike Haug forschen an einem unabhängigen, gemeinnützigen Forschungsinstitut, dem Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS. Das BIPS führt mitunter Studien durch, welche durch die Arzneimittelindustrie gefördert werden. Dabei handelt es sich fast ausschließlich um Sicherheitsstudien nach der Zulassung (PASS), die von Gesundheitsbehörden angefordert werden. Die Planung und Durchführung dieser Studien sowie deren Interpretation und Publikation werden nicht von der Arzneimittelindustrie beeinflusst. Die vorgestellte Studie wurde nicht von der Arzneimittelindustrie gefördert und wurde unter Einhaltung des ENCePP-Verhaltenskodex durchgeführt.

Danksagung

Die Autoren und Autorinnen danken den Gesetzlichen Krankenkassen AOK Bremen/Bremerhaven, Die Techniker (TK), DAK-Gesundheit und hkk Krankenkasse, die Daten für diese Studie bereitgestellt haben.

ORCIDs der Autoren und Autorinnen

- Asendorf, Marieke, B.A.:
<https://orcid.org/0000-0002-8918-3709>
- Reinold, Jonas, EU-M.Sc.:
<https://orcid.org/0000-0001-8266-2574>
- Dr. rer. medic. Schink, Tania:
<https://orcid.org/0000-0002-0224-1866>
- Dr. rer. nat. Kollhorst, Bianca:
<https://orcid.org/0000-0001-5964-954X>
- Prof. Dr. sc. hum. Haug, Ulrike:
<https://orcid.org/0000-0002-1886-2923>

Anhänge

Verfügbar unter <https://doi.org/10.3205/mibe000235>

1. Anhang1_mibe000235.pdf (439 KB)
Appendix

Literatur

1. Geyer S. Sozialstruktur und Krankheit. Analysen mit Daten der Gesetzlichen Krankenversicherung [Social inequalities in health. Analysis using data from statutory health insurance companies]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008 Oct;51(10):1164-72. DOI: 10.1007/s00103-008-0651-1
2. Richter M, Hurrelmann K. Gesundheitliche Ungleichheit: Ausgangsfragen und Herausforderungen. In: Richter M, Hurrelmann K, editors. Gesundheitliche Ungleichheit. Wiesbaden: VS Verlag für Sozialwissenschaften; 2009. p. 11-31.
3. Wilkinson R, Marmot M. Soziale Determinanten von Gesundheit. Die Fakten. 2. Ausgabe. Kopenhagen: World Health Organization Europe; 2004.
4. Greiner W, Batram M, Damm O, Scholz S, Witte J. Kinder- und Jugendreport 2018 – Gesundheitsvorsorge von Kindern und Jugendlichen in Deutschland. Schwerpunkt: Familiengesundheit. Bielefeld & Hamburg; 2018.
5. Lamerz A, Kuepper-Nybelen J, Wehle C, Bruning N, Trost-Brinkhues G, Brenner H, Hebebrand J, Herpertz-Dahlmann B. Social class, parental education, and obesity prevalence in a study of six-year-old children in Germany. *Int J Obes (Lond)*. 2005 Apr;29(4):373-80. DOI: 10.1038/sj.ijo.0802914
6. Lampert T, Kroll LE, von der Lippe E, Müters S, Stolzenberg H. Sozioökonomischer Status und Gesundheit: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) [Socioeconomic status and health: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2013 May;56(5-6):814-21. DOI: 10.1007/s00103-013-1695-4
7. Mensink GB, Schienkiewitz A, Haftenberger M, Lampert T, Ziese T, Scheidt-Nave C. Übergewicht und Adipositas in Deutschland: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) [Overweight and obesity in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2013 May;56(5-6):786-94. DOI: 10.1007/s00103-012-1656-3
8. OECD. Health at a Glance. Paris: OECD Publishing; 2019.
9. Großschädl F, Stronegger WJ. Long-term trends (1973-14) for obesity and educational inequalities among Austrian adults: men in the fast lane. *Eur J Public Health*. 2019 Aug;29(4):790-6. DOI: 10.1093/eurpub/cky280
10. Hajek A, Lehnert T, Ernst A, Lange C, Wiese B, Prokein J, Weyerer S, Werle J, Pentzek M, Fuchs A, Luck T, Bickel H, Mösch E, Hesel K, Wagner M, Maier W, Scherer M, Riedel-Heller SG, König HH; AgeCoDe Study Group. Prevalence and determinants of overweight and obesity in old age in Germany. *BMC Geriatr*. 2015 Jul;15:83. DOI: 10.1186/s12877-015-0081-5
11. Damm K, Lange A, Zeidler J, Braun S, Graf von der Schulenburg JM. Einführung des neuen Tätigkeitsschlüssels und seine Anwendung in GKV-Routinedatenauswertungen. Möglichkeiten und Limitationen [Implementation of the new German job role code and its application in claims data analysis. Possibilities and limitations]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2012 Feb;55(2):238-44. DOI: 10.1007/s00103-011-1418-7
12. Bundesagentur für Arbeit. Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Nürnberg: Bundesagentur für Arbeit; 2007.
13. Bundesagentur für Arbeit. Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe 2010. Stand: Mai 2017. Nürnberg: Bundesagentur für Arbeit; 2010.
14. Kroll LE, Schumann M, Hoebel J, Lampert T. Regionale Unterschiede in der Gesundheit – Entwicklung eines sozioökonomischen Deprivationsindex für Deutschland. *J Health Monit*. 2017;2(2):103-20. DOI: 10.17886/RKI-GBE-2017-035
15. Robert-Koch-Institut, Abteilung für Epidemiologie und Gesundheitsmonitoring. Gesundheit in Deutschland aktuell (GEDA 2014/2015-EHIS). Scientific Use File. 1. Version. Berlin: Robert-Koch-Institut; 2015.

16. Saß AC, Lange C, Finger JD, Allen J, Born S, Hoebel J, Kuhnert R, Müters S, Thelen J, Schmich P, Varga M, von der Lippe E, Wetzstein M, Ziese T. „Gesundheit in Deutschland aktuell“ – Neue Daten für Deutschland und Europa Hintergrund und Studienmethodik von GEDA 2014/2015-EHIS. *J Health Monit.* 2017;2(1):83-90. DOI: 10.17886/RKI-GBE-2017-012
17. Maukonen M, Männistö S, Tolonen H. A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review. *Scand J Public Health.* 2018 Jul;46(5):565-79. DOI: 10.1177/1403494818761971
18. Schubert I, Köster I, Küpper-Nybelen J, Ihle P. Versorgungsforschung mit GKV-Routinedaten. Nutzungsmöglichkeiten versichertenbezogener Krankenkassendaten für Fragestellungen der Versorgungsforschung [Health services research based on routine data generated by the SHI. Potential uses of health insurance fund data in health services research]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2008 Oct;51(10):1095-105. DOI: 10.1007/s00103-008-0644-0
19. Geyer S. Die Bestimmung der sozioökonomischen Position in Prozessdaten und ihre Verwendung in Sekundärdatenanalysen. In: Swart E, Ihle P, Gothe H, Matusiewicz D, editors. *Routinedaten im Gesundheitswesen Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven.* Bern: Verlag Hans Huber; 2005. p. 203-7.
20. Lidfeldt J, Li TY, Hu FB, Manson JE, Kawachi I. A prospective study of childhood and adult socioeconomic status and incidence of type 2 diabetes in women. *Am J Epidemiol.* 2007 Apr;165(8):882-9. DOI: 10.1093/aje/kwk078
21. Muschik D, Jaunzeme J, Geyer S. Are spouses' socio-economic classifications interchangeable? Examining the consequences of a commonly used practice in studies on social inequalities in health. *Int J Public Health.* 2015 Dec;60(8):953-60. DOI: 10.1007/s00038-015-0744-1
22. Lampert T, Kroll LE. Die Messung des sozioökonomischen Status in sozialepidemiologischen Studien. In: Richter M, Hurrelmann K, editors. *Gesundheitliche Ungleichheit: Grundlagen, Probleme, Perspektiven.* Wiesbaden: VS Verlag für Sozialwissenschaften; 2009. p. 309-34.
23. Jacobs E, Tönnies T, Rathmann W, Brinks R, Hoyer A. Association between regional deprivation and type 2 diabetes incidence in Germany. *BMJ Open Diabetes Res Care.* 2019;7(1):e000857. DOI: 10.1136/bmjdr-2019-000857
24. Kaulh B, Maier W, Schweikart J, Keste A, Moskwyn M. Exploring the small-scale spatial distribution of hypertension and its association to area deprivation based on health insurance claims in Northeastern Germany. *BMC Public Health.* 2018 Jan;18(1):121. DOI: 10.1186/s12889-017-5017-x
25. Grobe T, Ihle P. Versichertenstammdaten und sektorübergreifende Analyse. In: Swart E, Ihle P, Gothe H, Matusiewicz D, editors. *Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven.* Bern: Verlag Hans Huber; 2005. p. 17-34.
26. Wahrendorf M, Rupprecht CJ, Dortmann O, Scheider M, Dragano N. Erhöhtes Risiko eines COVID-19-bedingten Krankenhausaufenthaltes für Arbeitslose: Eine Analyse von Krankenkassendaten von 1,28 Mio. Versicherten in Deutschland [Higher risk of COVID-19 hospitalization for unemployed: an analysis of health insurance data from 1.28 million insured individuals in Germany]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2021 Mar;64(3):314-21. DOI: 10.1007/s00103-021-03280-6

Korrespondenzadresse:

Prof. Dr. sc. hum. Ulrike Haug
Leibniz-Institut für Präventionsforschung und
Epidemiologie – BIPS, Achterstraße 30, 28359 Bremen,
Deutschland
haug@leibniz-bips.de

Bitte zitieren als

Asendorf M, Reinold J, Schink T, Kollhorst B, Haug U. Abbildbarkeit des sozioökonomischen Status in der pharmakoepidemiologischen Forschungsdatenbank GePaRD: Beschreibung und Anwendung am Beispiel des Zusammenhangs mit Adipositas. *GMS Med Inform Biom Epidemiol.* 2022;18(1):Doc02.
DOI: 10.3205/mibe000235, URN: urn:nbn:de:0183-mibe0002350

Artikel online frei zugänglich unter

<https://doi.org/10.3205/mibe000235>

Veröffentlicht: 12.05.2022

Copyright

©2022 Asendorf et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.

Assessing the socioeconomic status in the German Pharmacoepidemiological Research Database (GePaRD): Description and exemplary application using the association with obesity

Abstract

Background: The socioeconomic status is often an important confounder in the context of epidemiological studies. However, this information is typically not included in studies based on claims data, possibly due to a lack of knowledge regarding the completeness and validity of the available information.

Objectives: Our aim was to assess and evaluate the socioeconomic status in the German Pharmacoepidemiological Research Database (GePaRD).

Methodology: First, an algorithm was developed to assign an educational status to individuals insured in 2017 in order to estimate the socioeconomic status at the individual level. For comparison, the socioeconomic status was estimated based on the place of residence. For plausibility, we examined whether the known association between the presence of obesity and the socioeconomic status could be reproduced.

Results: Based on different age groups it was possible to dichotomously estimate the socioeconomic status for 86–93% of those under 60 years of age and for 67% of the 60- to 69-year-olds. For individuals 70 years and older, the proportion of missing values was very high. In all subgroups, the prevalence of obesity was higher among persons of lower socioeconomic status than among persons of higher socioeconomic status. The estimation based on the place of residence produced plausible results, but with smaller differences by socioeconomic status.

Discussion and conclusion: Overall, the study results suggest that in terms of plausibility and completeness the information available in GePaRD is well suited for estimating the individual socioeconomic status for age groups up to 69 years.

Keywords: socioeconomic status, educational status, obesity, index of social deprivation, electronic health care data

Marieke Asendorf¹

Jonas Reinold¹

Tania Schink²

Bianca Kollhorst³

Ulrike Haug^{1,4}

1 Department of Clinical Epidemiology, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

2 Department of Clinical Epidemiology, Leibniz Institute for Prevention Research and Epidemiology – BIPS Bremen, Germany

3 Department of Biometry and Data Management, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

4 Department of Human and Health Sciences, University of Bremen, Germany

Introduction

For decades, scientific studies have shown a social imbalance in the distribution of the burden of various diseases. In most cases, socially disadvantaged groups bear the greater burden of disease [1], [2], [3]. For example, there is a pronounced association between the socioeconomic status and the presence of obesity [4], [5], [6], [7], [8]. Based on survey data, Großschädl/Stronegger report a 13-percentage point higher obesity prevalence for women with primary/vocational school education compared to those with university entrance qualification. For men, a difference of about 9 percentage points was observed in the study [9].

Consequently, it is important to integrate information on the socioeconomic status into epidemiological studies as a possible influencing or confounding variable. Studies based on primary data usually collect information on school education and vocational qualification, occupation or income for this purpose [6], [10]. In addition to primary data, health insurance claims data are an increasingly important data source for health services and epidemiological research. For example, the German Pharmacoepidemiological Research Database (GePaRD) contains claims data from four statutory health insurance providers, covers approximately 20% of the German population, and currently includes the data years 2004–2017.

GePaRD is utilized for numerous studies in the areas of drug utilization and drug safety research and cancer screening. In many of these studies, the incorporation of the socioeconomic status is also of great relevance or interest. However, it has not yet been systematically investigated to what extent variables available in GePaRD are suitable to characterize the socioeconomic status and to what degree of completeness they are available. There are several ways of estimating the socioeconomic status in data from statutory health insurance providers (see Table 1). In GePaRD, two variables could provide information on the socioeconomic status. First, the data include the county code of the residence of the insured, which – linked to the regional deprivation index of the respective county – may be used to estimate the socioeconomic status based on aggregated information, i.e., the same socioeconomic status is assigned to all persons in a county. Second, there is a variable containing information on occupation, occupational status, and education of the insured according to the key list of the Federal Employment Agency, thus providing a way to estimate the socioeconomic status on an individual level. The aim of this study was to investigate the completeness of these two types of information in GePaRD and to indirectly assess their informative value and plausibility with regard to the socioeconomic status by determining the association with the presence of obesity.

Methods

Data source and study population

We used the German Pharmacoepidemiological Research Database (GePaRD) for this study. GePaRD is based on claims data from four statutory health insurance providers in Germany and currently includes information on approximately 25 million persons who have been insured with one of the participating providers since 2004 or later. In addition to demographic data, GePaRD contains information on drug dispensings as well as outpatient and inpatient services and diagnoses. Per data year, there is information on approximately 20% of the general German population and all geographical regions of Germany are represented. For this study, we used data from the years 2004 to 2017. We included all persons who were insured for at least one day in 2017 and were alive on that day, for whom information on sex, year of birth, and place of residence was available and whose place of residence was in Germany.

Estimation of the socioeconomic status on the individual level

The starting point was a variable included in the master data of insured persons containing information on occupation, occupational status, and education. This information, summarized by the so-called occupation key (German: Tätigkeitsschlüssel) is transmitted at least annually

by employers to the social insurance institutions for all persons covered by statutory health insurance and for primary insurance holders covered voluntarily. The transfer occurs in accordance with Section 28a of the German Social Code (German: Sozialgesetzbuch – SGB) IV. This information is thus not available for persons who have retired in the respective year or for “other” insured persons and co-insured family members [11].

For the operationalization of this variable, the following principles were applied:

A) The focus was to estimate the socioeconomic status based on information regarding school education.

B) Where unambiguous, information on occupational degree or status was included to draw conclusions about school education. For example, individuals with a university degree were assumed to have a (technical) degree allowing university entrance regardless of the information on their school education.

C) In order to assign an educational status to persons insured in 2017, earlier data years were also used. For example, in the case of some retired persons, this made it possible to identify past periods of employment, i.e., years in which an occupation key had still been transmitted.

D) Individuals covered by the insurance of family members were assumed to have the same socioeconomic status as the associated primary insurance holders. Based on this assumption, we attempted to link individuals for whom an occupation code was not available in any of the data years to a family member with joint insurance in GePaRD to whom a socioeconomic status had already been assigned. The family-ID contained in the master data of the insured was used for this linkage.

E) If data from different calendar years yielded discrepant information on education, the highest level of education was always used. This is based on the assumption that the level of education acquired does not decrease, but may at best increase as a result of further school education.

Based on this operationalization, insured persons were classified into the following three categories:

1. (technical) university entrance qualification (German: (fachgebundene) Hochschulreife),
2. basic secondary degree or secondary degree (German: Hauptschulabschluss oder mittlere Reife),
3. school education unknown.

In the last category, insured persons with “school education unknown” or “no degree” and those without data on school education were grouped together, since it was expected that the group with no degree would be small. This classification is based on the options provided by the occupation key before it was changed to a new classification system in 2011 [11], [12], [13]. The new classification system offers the additional option of distinguishing between the basic secondary degree and the secondary degree (see Figure 1). However, for the most part, the occupation key corresponding to the new classification was transmitted to GePaRD only starting with the 2017

Table 1: Overview of information for the estimation of the socioeconomic status in data from statutory health insurance providers

Characteristic	Possible limitations	Availability in GePaRD
Premium group		
Defines the assignment to different premium groups [25]		No
Group of insured persons as a combination of type of insured person and person group key		
Facilitates, among other things, a distinction between voluntarily insured employees, individuals in low-paid employment with social benefits, and pension recipients [26]		No
Occupation		
Defines the occupation performed according to the classification of occupations of the Federal Employment Agency	Available only during periods of employment (primary insurance holder)	Yes
General school qualification		
Defines the highest general school qualification	Available only during periods of employment (primary insurance holder)	Yes
Vocational training qualification		
Defines the highest vocational training qualification	Available only during periods of employment (primary insurance holder)	Yes
Remuneration		
Specifies the income for the calculation of the premium [25]	Available only up to a certain income threshold; available only for employees subject to social insurance contributions	No
Place of residence with zip code		
Can be used in combination with regional indicators [25]		Available at the county level

data set. To investigate the potential of the new occupation key as well, we conducted sensitivity analyses. For this purpose, the population of the main analysis was restricted to persons with an educational status based on information on vocational or school education of the classification system after 2011.

Estimation of the socioeconomic status based on information on the place of residence

The most recent information on residence at the county level available in GePaRD was linked to the German Index of Socioeconomic Deprivation (GISD) developed by the Robert Koch Institute (RKI) in its revised version based on 2014 data. The index reflects the extent of deprivation based on the three dimensions income, education, and occupation [14]. In line with the smallest spatial unit in GePaRD, the GISD was used at the county level (see appendix (Attachment 1) for further explanation). Analogously to the procedure of Kroll et al., the deprivation scores assigned to counties were divided into quintiles (first quintile: “low deprivation”; second to fourth quintile: “medium deprivation”; fifth quintile: “high deprivation” [14].

Collection of information on the presence of obesity

The presence of obesity was determined on the basis of the relevant ICD-10 codes (inpatient or confirmed outpatient) and codes for procedures indicating obesity therapy (e.g., bariatric surgery). Persons with at least one of the relevant codes in GePaRD were classified as obese. For these codes, we considered not only the data year 2017, but also all data years prior to 2017 available in GePaRD for each person given that obesity is generally a chronic and therefore permanent health condition. Since it is presumably often only coded if relevant as a reason for billing, which may not be the case every year, reverting to earlier data years increases the sensitivity of detecting obesity.

Analysis of the data in GePaRD

First, to assess the completeness of individual schooling information in GePaRD, we determined the proportion of the insured in 2017 to whom school education could be assigned based on the algorithm described above. To determine the prevalence of obesity, the persons insured in 2017 who were classified as obese as described above were considered in the numerator and all persons from the described cohort of the insured were considered in

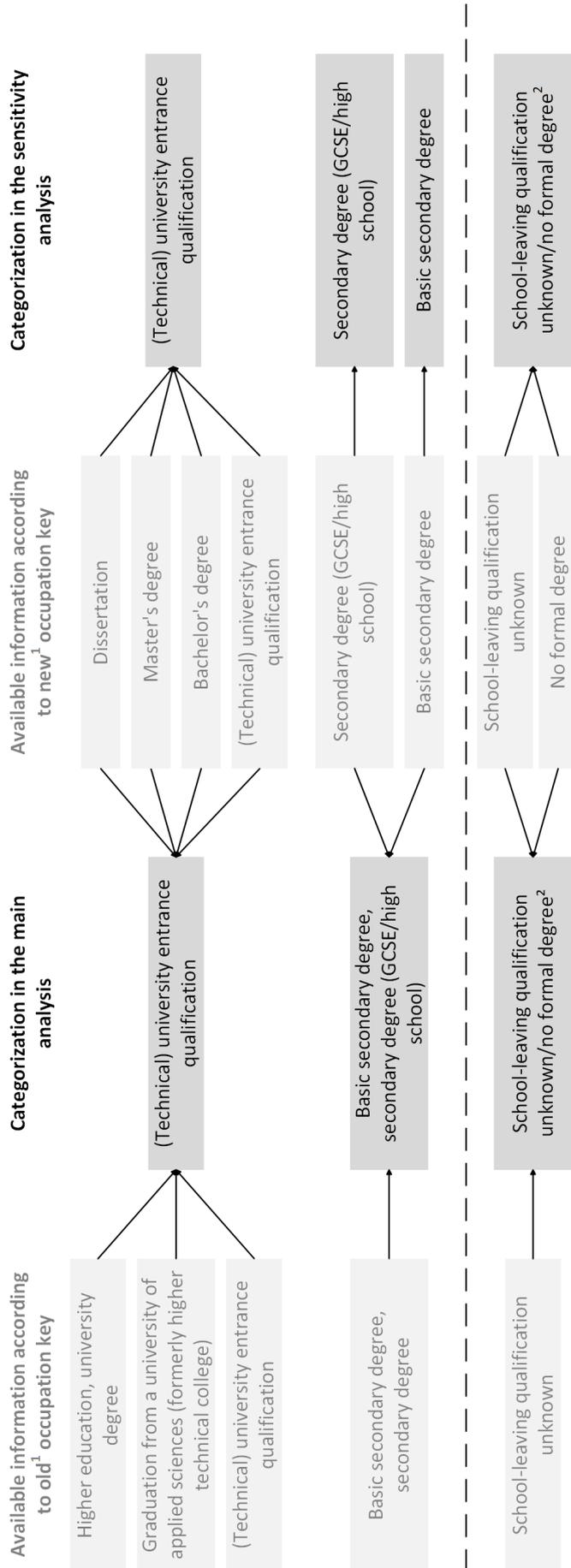


Figure 1: Categorization of information on education in GePaRD

¹The term "old occupation key" refers to the occupation key before the 2011 transition by the Federal Employment Agency. The term "new occupation key" refers to the occupation key after the transition. In GePaRD, data on the new occupation key were supplied (also retrospectively) for the first time with the data delivery for 2017 for all health insurance providers.

²The category "school-leaving qualification unknown/no formal degree" was not included in analyses on the association with obesity prevalence.

the denominator. The obesity prevalence calculated in this way was determined in the main analyses stratified by (technical) university entrance qualification vs. basic secondary degree or secondary degree (dichotomous classification). In the sensitivity analysis, prevalence was stratified according to the three categories (technical) university entrance qualification, secondary degree, and basic secondary degree. For the analyses based on the GISD, we followed an analogous procedure, i.e., obesity prevalence was determined stratified according to the categories low, medium, and high socioeconomic deprivation. All analyses were stratified by age and sex of the insured.

Comparative analysis of primary data

In addition to the analyses based on GePaRD, primary data collected as part of the GEDA 2014/2015 survey of the RKI were comparatively analyzed [15], i.e., obesity prevalence was also determined according to age, sex, and school education (in the categories (technical) university entrance qualification, secondary degree, and basic secondary degree). The survey, conducted by means of electronic questionnaires or telephone interviews, involved 24,016 persons aged 18 years and older with main residence in Germany [16].

Results

After applying the inclusion and exclusion criteria, 17,317,559 individuals were included in the study (proportion of women: 53%). 258,013 individuals were excluded due to missing information (in 84% of these individuals, information on the place of residence in Germany was missing or not valid). Overall, the socioeconomic status could be approximated at the individual level according to the algorithm described for 73% of women and 78% of men.

For about half of the men and women, there was information on their educational level. Additionally, for about 20%, there was no information on their education, but it was possible to match them to family members with information on their education and thus indirectly estimate their socioeconomic status (see Figure 2). The proportion of individuals to whom a socioeconomic status could be assigned with the algorithm was consistently above 85% for men and women in the age groups under 60, it was 67% for 60- to 69-year-olds (women: 63%, men: 72%), 16% for 70- to 79-year-olds (women: 10%, men: 23%), and for persons aged 80 and over below 1% (women: 0.3%, men: 1%) (see Table 2). The proportion of individuals without a formal degree (no school completed) among all individuals was less than 1% across all subgroups.

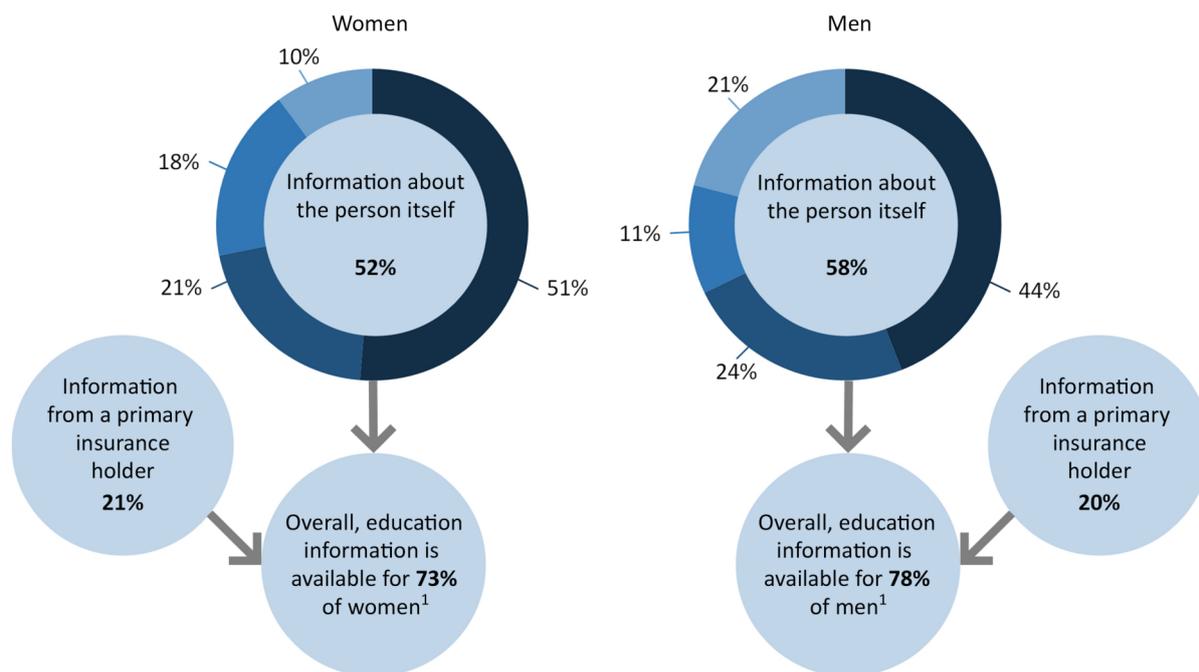
Figure 3 shows the prevalence of obesity observed in GePaRD stratified by age group, sex, and assigned educational level (basic secondary degree/secondary degree and (technical) university entrance qualification) at the individual level. In all age groups and in both sexes, the

prevalence of obesity, which increased overall with age, was higher in individuals with lower educational level. For example, the point estimates of prevalence among women aged 30–39 years were approximately 9 percentage points higher than of those with higher education (22% vs. 13%). In the same age group, the difference for men was about 5 percentage points (12% vs. 7%). Fig. 7 (Attachment 1) illustrates the comparative obesity prevalence for all categories including education unknown, no formal education, and persons without any information on educational level. Fig. 8 (Attachment 1) also compares obesity prevalence for the two groups with and without information on education. This shows a largely consistent prevalence in both groups for men, except for 18- to 29-year-olds (lower obesity prevalence in the group without information on education). For women up to the age of 50, a pattern similar to that of men emerges. Older women show a higher obesity prevalence in the group with information on education than in women without such information.

In the sensitivity analyses classifying the educational status into three categories (basic secondary degree, secondary degree, and (technical) university entrance qualification), 27% of the insured could be included (24% of men, 29% of women). There was an inverse relationship between obesity prevalence and educational level in all subgroups under 80, with a smaller difference between the categories secondary degree vs. basic secondary degree (e.g., 5% percentage points among 30–39-year-old women) than between the categories secondary degree vs. (technical) university entrance qualification (8 percentage points among 30- to 39-year-old women) (Figure 4). The maximum difference in obesity prevalence between persons with a basic secondary school diploma vs. persons with a (technical) university entrance qualification was 13 percentage points and was observed in the subgroup of 30- to 39-year-old women.

The results on the association between the obesity prevalence observed in GePaRD and the socioeconomic status estimated by means of the spatial deprivation index of the place of residence (county level) are shown in Figure 5. Again, there is an increase in obesity prevalence with increasing deprivation in all age groups and for both sexes. The differences were less pronounced compared to the analysis based on the individually estimated socioeconomic status. The difference between the subgroups with the highest and lowest deprivation, respectively, is about 8 percentage points among 30–39-year-old women (21% vs. 13%) and about 4 percentage points among men in this age group (11% vs. 7%).

Figure 6 illustrates the results of the comparative analysis based on the GEDA 2014/2015 dataset. This dataset contains information on educational attainment at the individual level, i.e., the results can be viewed in comparison with the GePaRD results shown in Figure 3. Compared with GePaRD, some subgroups – particularly those in older age groups – showed somewhat different, often lower, obesity prevalences. However, overall, the patterns were similar in terms of differences between educational



Distribution of specific data sources on the individual level (the proportions of persons of each sex with educational information from the corresponding source are provided as relative values).

- Schooling (based on the new occupation key)
- Academic vocational qualification
- Occupational status
- Scholing (based on the old occupation key)

¹In this context, the code “no school-leaving qualification” is classified as a valid code and is included in this number accordingly.

Figure 2: Sources of information on education

Table 2: Proportion of insured persons to whom a socioeconomic status could be assigned by age and sex (in this context, the code “no school-leaving qualification” is classified as a valid code and is included in the calculated amounts accordingly)

Age	Assignment of a socioeconomic status		
	Both sexes	Women	Men
Younger than 18 years	88.4%	88.5%	88.4%
18–29 years	88.2%	88.7%	87.7%
30–39 years	92.6%	93.3%	92.0%
40–49 years	89.5%	90.3%	88.6%
50–59 years	85.7%	86.1%	85.3%
60–69 years	67.0%	63.4%	71.6%
70–79 years	15.7%	10.0%	23.3%
80 years and older	0.7%	0.3%	1.2%

groups. In tendency, the absolute difference in obesity prevalence between education groups was more pronounced for GEDA and showed fewer sex differences than for GePaRD. For example, among women aged 40–49, the difference between the higher and lower education groups was 7 percentage points (GePaRD:

7 percentage points) and among men aged 40–49, the difference was 8 percentage points (GePaRD: 4 percentage points).

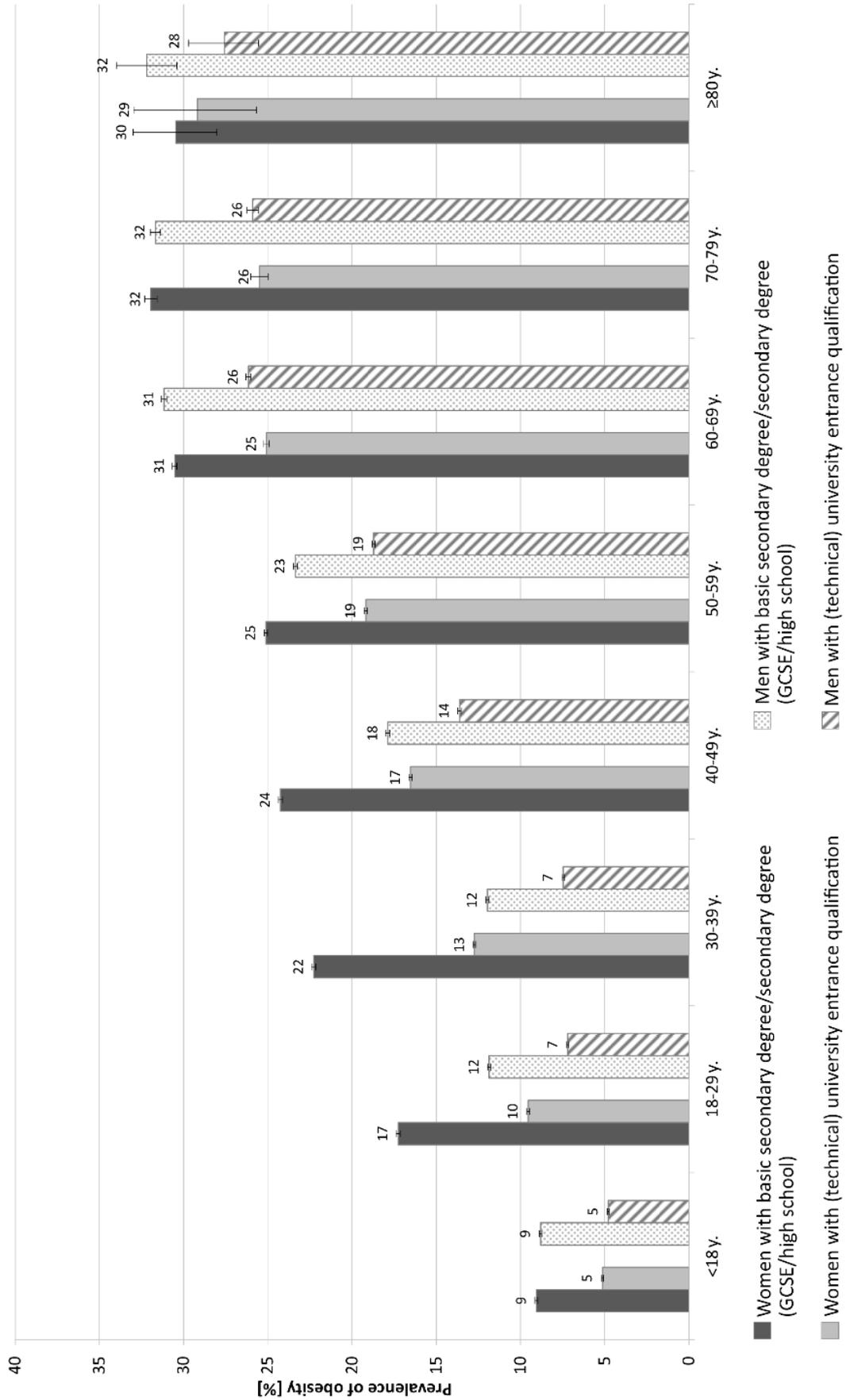


Figure 3: Prevalence of obesity in GePaRD for 2017 stratified by age, sex, and individually assigned educational level

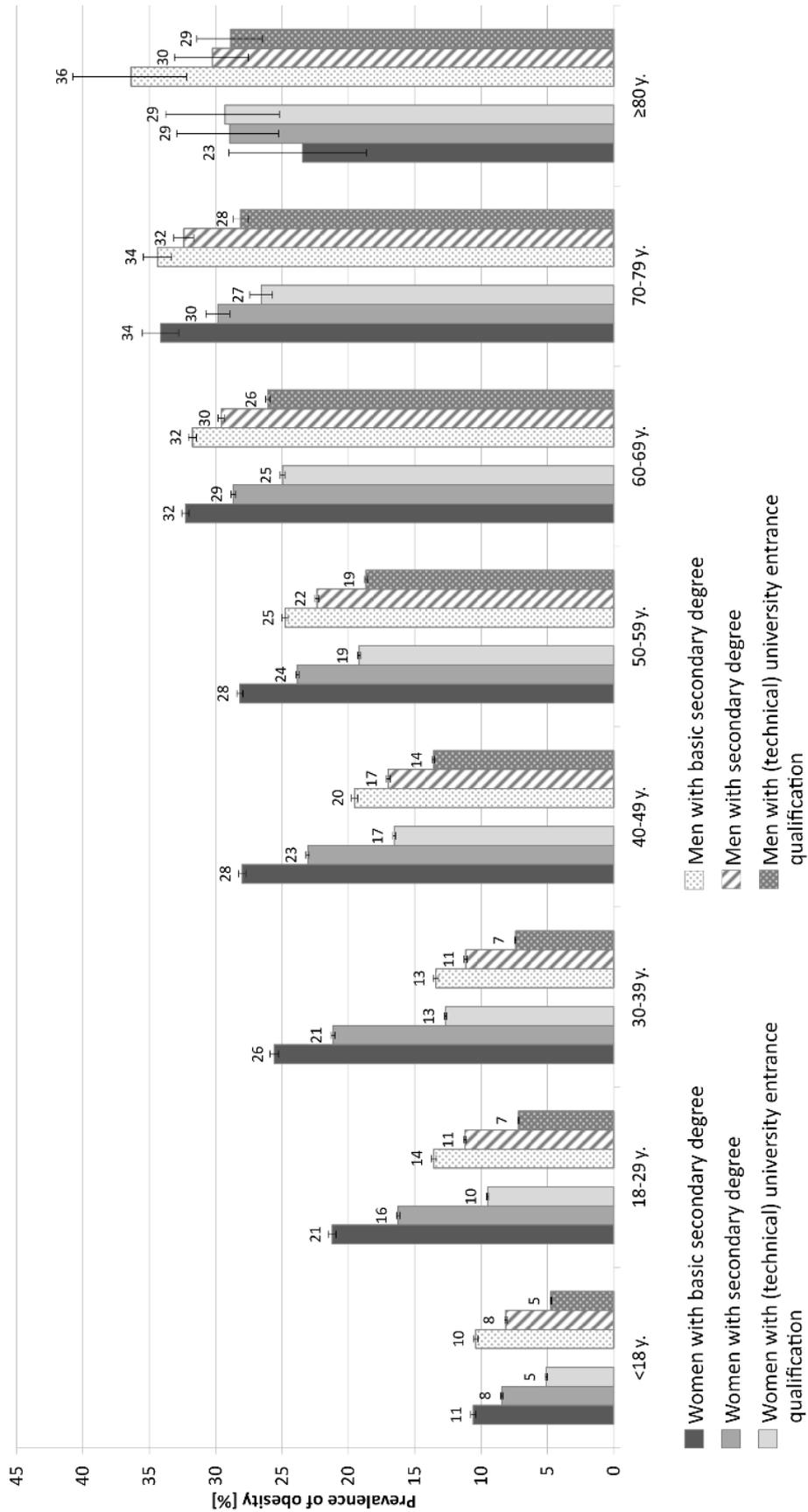


Figure 4: Prevalence of obesity in GePaRD for 2017 stratified by age, sex, and individually assigned level of education (in three categories based on new occupational key)

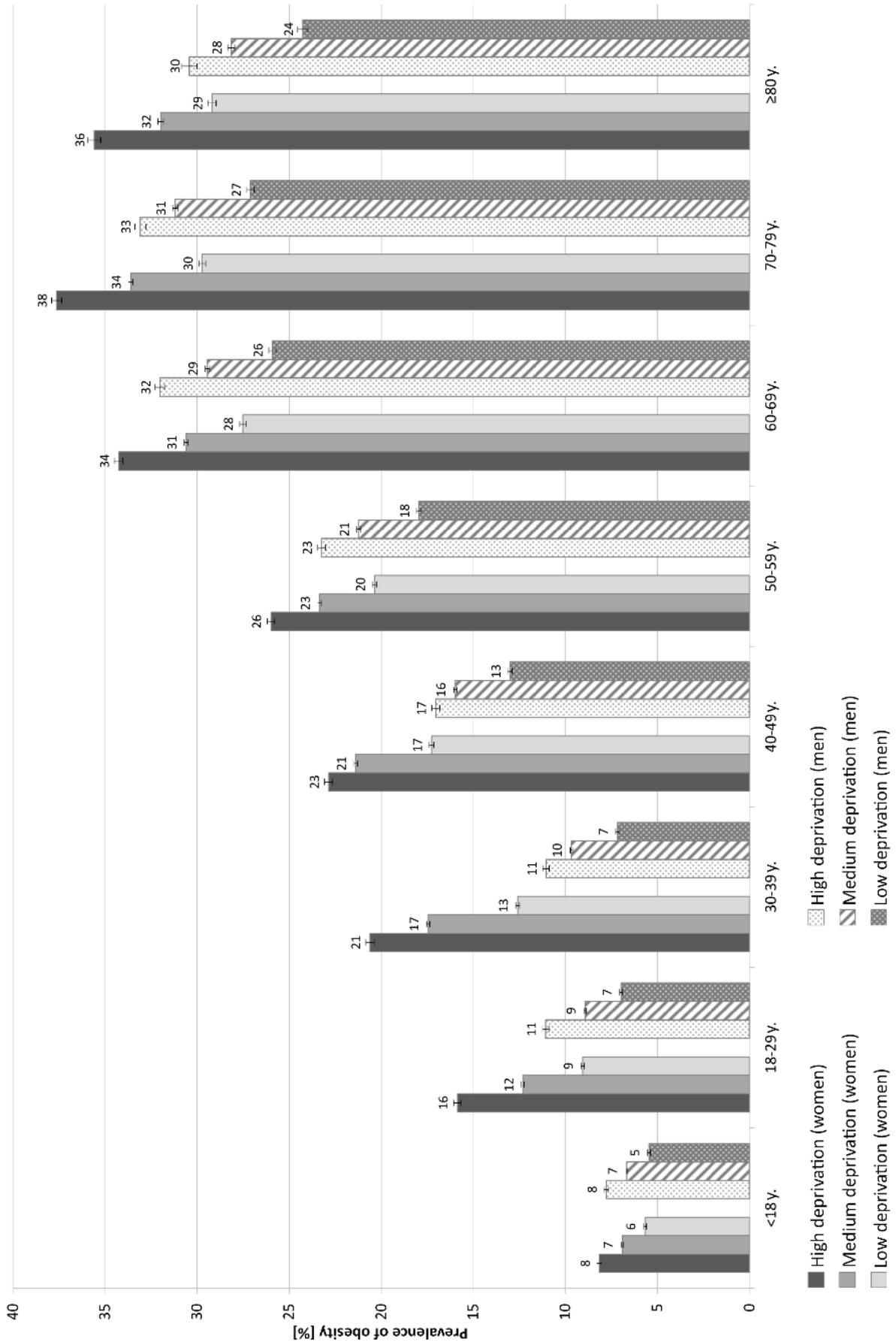


Figure 5: Prevalence of obesity in GePaRD for 2017 stratified by age, sex, and regional deprivation level (GISD)

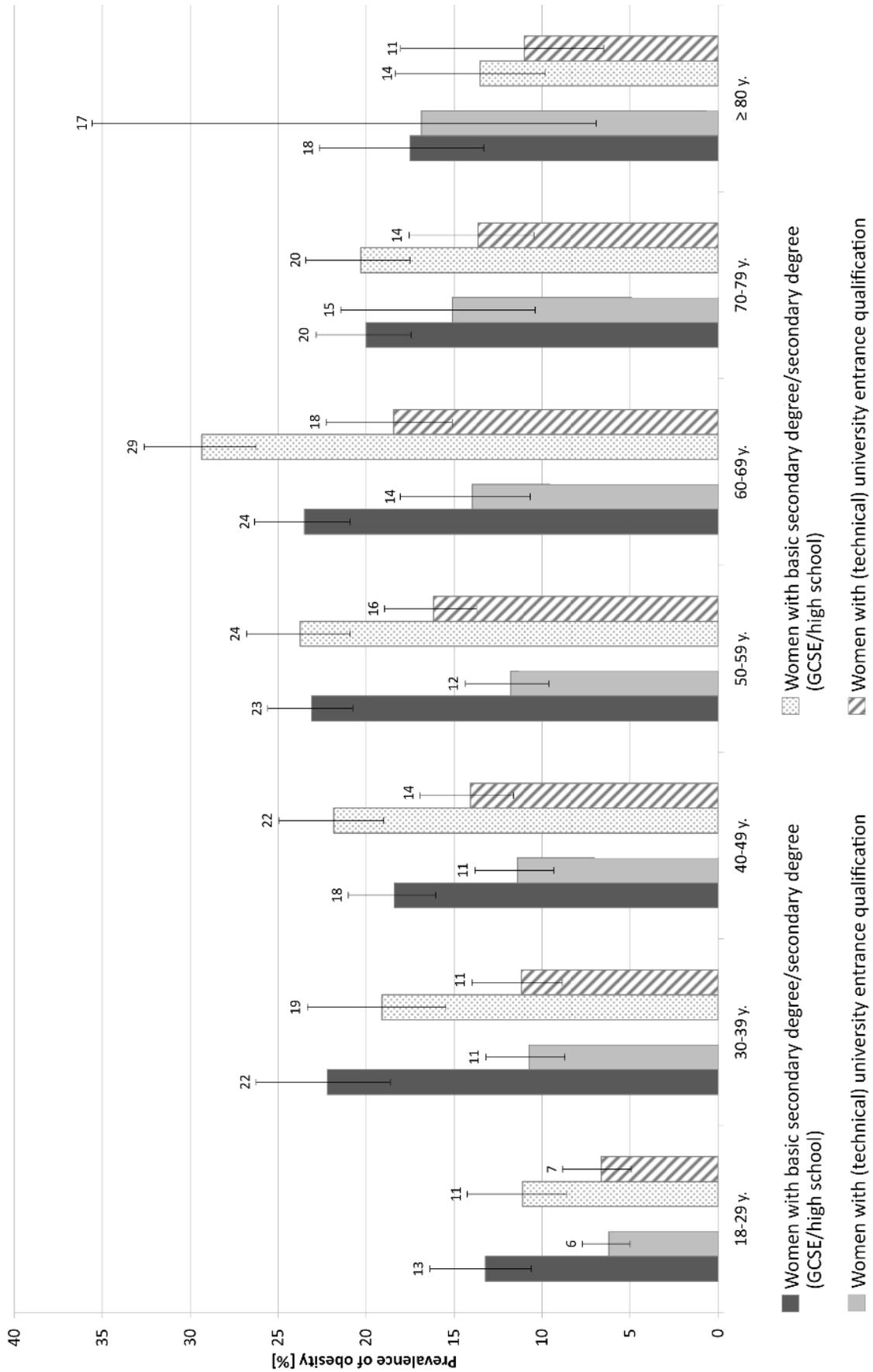


Figure 6: Prevalence of obesity in GEDA for 2014–2015 stratified by age, sex, and individually assigned educational level

Discussion

The results of this study show that based on individual information available in the Pharmacoepidemiological Research Database GePaRD, it is possible to classify about three quarters of the insured into a higher or lower educational category as an approximation of the socioeconomic status. The plausibility of this classification was confirmed by the reproduction of the known association between obesity and socioeconomic status based on this classification. Overall, the estimation of the socioeconomic status based on the county of residence also rendered plausible results. However, the difference observed in these analyses between 20% of the insured with the lowest and the highest deprivation score, respectively (i.e., 60% of the insured were not considered in the comparison), was similar in magnitude to the difference observed in the individual estimation of the socioeconomic status based solely on a dichotomous classification of all insured with available information on education. The differences were even more pronounced when the educational level was divided into three categories, which also indicates that the discriminatory power of the estimated socioeconomic status based on individual information is greater than in the case of an estimate based on the county of residence.

The magnitude of the association between the presence of obesity and the socioeconomic status, as well as the differences by age and sex, were overall consistent with our results of the comparative analyses based on GEDA data. For some subgroups, the difference by socioeconomic status was slightly more pronounced in the GEDA than in the GePaRD analyses. For example, among those aged 50–59 years, the GEDA analyses showed a 2-fold (women) and 1.5-fold (men) increase in obesity prevalence for those with a basic secondary degree or secondary degree compared with the higher education group, whereas the GePaRD analyses showed a 1.4-fold increase in prevalence for men and women. However, when comparing the results one should bear in mind that some of the estimators in the GEDA analyses have wide confidence intervals. It should also be noted that full agreement was not to be expected simply because of the different ways in which obesity was assessed. In insurance data, diagnoses appear primarily if relevant for billing purposes, i.e., if services were provided in connection with the diagnosis. A sensitive case definition was applied in the GePaRD analyses, which may have led to an overestimation of prevalence in adulthood. In the GEDA definition of obesity, which is based on self-reported height and weight, there may have been misclassification for other reasons. For example, it is known that body weight and height are often misreported in surveys [17]. According to Maukonen et al., this misclassification leads to an underestimation of obesity prevalence by 0.6 to 8.4 percentage points in European studies [17]. Thus, none of the data sources contains gold standard information on the presence of obesity, and differences in the strength of the association should therefore not be over-

interpreted. This generally also applies to the comparison with other studies, which, however, show similar overall associations between the individually assessed socioeconomic status and the presence of obesity [6], [7], [9].

The estimation of the individual socioeconomic status in our analyses was based on information related to the occupation key available in GePaRD. In principle, insurance data contain information that would facilitate an even more detailed estimate of the socioeconomic status, such as the premium level, which is indicative of income, or the exact place of residence. However, since this is very sensitive information, it is generally not available for secondary data use outside the health insurance provider. Most of the previous studies based on health insurance data taking the socioeconomic status into account therefore used (rough) information on the place of residence. An older study by Geyer based on data of 416,000 insured persons from the insurance provider AOK Mettmann from 1987 to 1996 already illustrated the potential of using information on education and occupation to estimate the individual socioeconomic status in insurance data [1], but this was given little consideration in more recent studies. On the one hand, this could be because the data analysts did not have access to the occupation key. On the other hand, it is also widely believed that the information is not available with the necessary completeness or unambiguity [18]. We were able to significantly increase completeness in the GePaRD analyses by exploiting the long lookback period in combination with the linkage of family members to primary insurance holders. Furthermore, we made assumptions, for example, that the socioeconomic status may be transferred from the primary insurance holder to family members. While this assumption is controversial in the literature [19], [20], [21], we believe it is justified and similar in terms of underlying assumptions to the use of family income in studies based on primary data.

With the GePaRD data source in combination with the methodology applied here, it was possible to roughly estimate the individual socioeconomic status (basic secondary degree/secondary degree and (technical) university entrance qualification) for 86–93% of those under 60. Even among 60- to 69-year-olds, this proportion was still 67% (men: 72% women: 63%). Undoubtedly and for several reasons, this is not a perfect classification of the socioeconomic status. Apart from the underlying assumptions, one cannot generally presume that the occupation keys reported by employers are completely free of errors [11]. Moreover, the educational level does not cover all dimensions of the socioeconomic status [22]. However, this rough estimate of the socioeconomic status is still of great value for numerous descriptive or analytical questions based on insurance data. This is particularly relevant since it is also an important surrogate parameter for other factors poorly represented in insurance data (e.g., lifestyle factors). For example, the socioeconomic status estimated in this way can be used for stratification and adjustment in order to identify or account for its relevance as an influencing or confounding factor, at least

to some extent. When considering the limitations mentioned above, it should be noted that estimating the socioeconomic status based on primary data is often also suboptimal due to missing information, among other reasons.

When using the GISD to estimate the socioeconomic status, the known association between low socioeconomic status (or high deprivation) and obesity was weaker than when the socioeconomic status was estimated individually in three categories, which was not surprising [23], [24]. Since the place of residence in GePaRD is available at the county level, we also used the GISD with this spatial reference. This results in 401 different regional units for Germany. This classification is very rough and does not take into account the differences in socioeconomic status between individuals in a region, i.e., considerable misclassification must be assumed. Accordingly, the estimation of the individual socioeconomic status based on GePaRD is preferable to the GISD, both dichotomously and in three categories. The GISD could be considered as a second-best option if the individual socioeconomic status cannot be estimated, which is currently often the case for individuals aged 70 and older, as it is often impossible to look back at the period of employment for them. However, in the long term, i.e., with an even longer lookback period in GePaRD, the proportion with missing occupation keys will likely continue to decrease among the elderly as well. Similarly, it is to be expected that educational and occupational information according to the new classification system will be available for ever more individuals, i.e., in the future, it will be possible to classify young people in particular into three educational categories.

The comparison of obesity prevalence among individuals with vs. without information on education (Fig. 8 in Attachment 1) shows differences between the two groups, particularly among 18- to 29-year-old men and women and among women aged 50 and older. In these groups, there appears to be a correlation between the presence of information on education and social status. In the case of 18- to 29-year-olds, this seems plausible as this group presumably includes students (higher educational level) who are still insured through their parents. Since some of the parents have already reached retirement age by that time resulting in missing occupation keys, there are also missing values for the co-insured family members. In contrast, persons undergoing vocational training (generally lower educational level) already earn a salary at this age and are self-insured, i.e., an occupation key is available. The pattern among older women may be explained by the fact that women with lower educational level are more often not employed and therefore co-insured with their husband, who is already retired (i.e., often lack an occupation key). Due to these correlations, the category “missing” should be considered separately in future analyses, especially with regard to 18- to 29-year-olds and women over 50.

Conclusion

Overall, the study results presented here indicate that the information available in GePaRD is well suited for the dichotomous estimation of the individual socioeconomic status for age groups up to 69 years, both in terms of completeness and plausibility. For individuals aged 70 and older, an estimation of the socioeconomic status based on the GISD could be used as a second-best option until a higher completeness can be achieved for these age groups in GePaRD as well – provided an even longer lookback period becomes available.

Notes

Acknowledgement

The authors would like to thank all statutory health insurance providers, which provided data for this study, namely AOK Bremen/Bremerhaven, DAK-Gesundheit, Die Techniker (TK), and hkk Krankenkasse.

Competing interests

Marieke Asendorf, Jonas Reinold, Tania Schink, Bianca Kollhorst, and Ulrike Haug are working at an independent, non-profit research institute, the Leibniz Institute for Prevention Research and Epidemiology – BIPS. Unrelated to this study, BIPS occasionally conducts studies financed by the pharmaceutical industry. Almost exclusively, these are post-authorization safety studies (PASS) requested by health authorities. The design and conduct of these studies as well as the interpretation and publication are not influenced by the pharmaceutical industry. The study presented was not funded by the pharmaceutical industry and was carried out in compliance with the ENCePP Code of Conduct.

The authors declare that they have no conflicts of interest related to this article.

ORCIDiDs of the authors

- Asendorf, Marieke, B.A.:
<https://orcid.org/0000-0002-8918-3709>
- Reinold, Jonas, EU-M.Sc.:
<https://orcid.org/0000-0001-8266-2574>
- Dr. rer. medic. Schink, Tania:
<https://orcid.org/0000-0002-0224-1866>
- Dr. rer. nat. Kollhorst, Bianca:
<https://orcid.org/0000-0001-5964-954X>
- Prof. Dr. sc. hum. Haug, Ulrike:
<https://orcid.org/0000-0002-1886-2923>

Attachments

Available from <https://doi.org/10.3205/mibe000235>

1. Attachment1_mibe000235.pdf (653 KB)
Appendix

References

1. Geyer S. Sozialstruktur und Krankheit. Analysen mit Daten der Gesetzlichen Krankenversicherung [Social inequalities in health. Analysis using data from statutory health insurance companies]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008 Oct;51(10):1164-72. DOI: 10.1007/s00103-008-0651-1
2. Richter M, Hurrelmann K. Gesundheitliche Ungleichheit: Ausgangsfragen und Herausforderungen. In: Richter M, Hurrelmann K, editors. Gesundheitliche Ungleichheit. Wiesbaden: VS Verlag für Sozialwissenschaften; 2009. p. 11-31.
3. Wilkinson R, Marmot M. Soziale Determinanten von Gesundheit. Die Fakten. 2. Ausgabe. Kopenhagen: World Health Organization Europe; 2004.
4. Greiner W, Batram M, Damm O, Scholz S, Witte J. Kinder- und Jugendreport 2018 – Gesundheitsvorsorge von Kindern und Jugendlichen in Deutschland. Schwerpunkt: Familiengesundheit. Bielefeld & Hamburg; 2018.
5. Lamerz A, Kuepper-Nybelen J, Wehle C, Bruning N, Trost-Brinkhues G, Brenner H, Hebebrand J, Herpertz-Dahlmann B. Social class, parental education, and obesity prevalence in a study of six-year-old children in Germany. Int J Obes (Lond). 2005 Apr;29(4):373-80. DOI: 10.1038/sj.ijo.0802914
6. Lampert T, Kroll LE, von der Lippe E, Müters S, Stolzenberg H. Sozioökonomischer Status und Gesundheit: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) [Socioeconomic status and health: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2013 May;56(5-6):814-21. DOI: 10.1007/s00103-013-1695-4
7. Mensink GB, Schienkiewitz A, Haftenberger M, Lampert T, Ziese T, Scheidt-Nave C. Übergewicht und Adipositas in Deutschland: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) [Overweight and obesity in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2013 May;56(5-6):786-94. DOI: 10.1007/s00103-012-1656-3
8. OECD. Health at a Glance. Paris: OECD Publishing; 2019.
9. Großschädl F, Stronegger WJ. Long-term trends (1973-14) for obesity and educational inequalities among Austrian adults: men in the fast lane. Eur J Public Health. 2019 Aug;29(4):790-6. DOI: 10.1093/eurpub/cky280
10. Hajek A, Lehnert T, Ernst A, Lange C, Wiese B, Prokein J, Weyerer S, Werle J, Pentzek M, Fuchs A, Luck T, Bickel H, Mösch E, Hesel K, Wagner M, Maier W, Scherer M, Riedel-Heller SG, König HH; AgeCoDe Study Group. Prevalence and determinants of overweight and obesity in old age in Germany. BMC Geriatr. 2015 Jul;15:83. DOI: 10.1186/s12877-015-0081-5
11. Damm K, Lange A, Zeidler J, Braun S, Graf von der Schulenburg JM. Einführung des neuen Tätigkeitsschlüssels und seine Anwendung in GKV-Routinedatenauswertungen. Möglichkeiten und Limitationen [Implementation of the new German job role code and its application in claims data analysis. Possibilities and limitations]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2012 Feb;55(2):238-44. DOI: 10.1007/s00103-011-1418-7
12. Bundesagentur für Arbeit. Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Nürnberg: Bundesagentur für Arbeit; 2007.
13. Bundesagentur für Arbeit. Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe 2010. Stand: Mai 2017. Nürnberg: Bundesagentur für Arbeit; 2010.
14. Kroll LE, Schumann M, Hoebel J, Lampert T. Regionale Unterschiede in der Gesundheit – Entwicklung eines sozioökonomischen Deprivationsindex für Deutschland. J Health Monit. 2017;2(2):103-20. DOI: 10.17886/RKI-GBE-2017-035
15. Robert-Koch-Institut, Abteilung für Epidemiologie und Gesundheitsmonitoring. Gesundheit in Deutschland aktuell (GEDA 2014/2015-EHIS). Scientific Use File. 1. Version. Berlin: Robert-Koch-Institut; 2015.
16. Saß AC, Lange C, Finger JD, Allen J, Born S, Hoebel J, Kuhnert R, Müters S, Thelen J, Schmich P, Varga M, von der Lippe E, Wetzstein M, Ziese T. „Gesundheit in Deutschland aktuell“ – Neue Daten für Deutschland und Europa Hintergrund und Studienmethodik von GEDA 2014/2015-EHIS. J Health Monit. 2017;2(1):83-90. DOI: 10.17886/RKI-GBE-2017-012
17. Maukonen M, Männistö S, Tolonen H. A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review. Scand J Public Health. 2018 Jul;46(5):565-79. DOI: 10.1177/1403494818761971
18. Schubert I, Köster I, Kupper-Nybelen J, Ihle P. Versorgungsforschung mit GKV-Routinedaten. Nutzungsmöglichkeiten versichertenbezogener Krankenkassendaten für Fragestellungen der Versorgungsforschung [Health services research based on routine data generated by the SHI. Potential uses of health insurance fund data in health services research]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008 Oct;51(10):1095-105. DOI: 10.1007/s00103-008-0644-0
19. Geyer S. Die Bestimmung der sozioökonomischen Position in Prozessdaten und ihre Verwendung in Sekundärdatenanalysen. In: Swart E, Ihle P, Gothe H, Matusiewicz D, editors. Routinedaten im Gesundheitswesen Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. Bern: Verlag Hans Huber; 2005. p. 203-7.
20. Lidfeldt J, Li TY, Hu FB, Manson JE, Kawachi I. A prospective study of childhood and adult socioeconomic status and incidence of type 2 diabetes in women. Am J Epidemiol. 2007 Apr;165(8):882-9. DOI: 10.1093/aje/kwk078
21. Muschik D, Jaunzeme J, Geyer S. Are spouses' socio-economic classifications interchangeable? Examining the consequences of a commonly used practice in studies on social inequalities in health. Int J Public Health. 2015 Dec;60(8):953-60. DOI: 10.1007/s00038-015-0744-1
22. Lampert T, Kroll LE. Die Messung des sozioökonomischen Status in sozialepidemiologischen Studien. In: Richter M, Hurrelmann K, editors. Gesundheitliche Ungleichheit: Grundlagen, Probleme, Perspektiven. Wiesbaden: VS Verlag für Sozialwissenschaften; 2009. p. 309-34.
23. Jacobs E, Tönnies T, Rathmann W, Brinks R, Hoyer A. Association between regional deprivation and type 2 diabetes incidence in Germany. BMJ Open Diabetes Res Care. 2019;7(1):e000857. DOI: 10.1136/bmjdr-2019-000857

24. KauhI B, Maier W, Schweikart J, Keste A, Moskwyn M. Exploring the small-scale spatial distribution of hypertension and its association to area deprivation based on health insurance claims in Northeastern Germany. *BMC Public Health*. 2018 Jan;18(1):121. DOI: 10.1186/s12889-017-5017-x
25. Grobe T, Ihle P. Versichertenstammdaten und sektorübergreifende Analyse. In: Swart E, Ihle P, Gothe H, Matusiewicz D, editors. *Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Bern: Verlag Hans Huber; 2005. p. 17-34.
26. Wahrendorf M, Rupprecht CJ, Dortmann O, Scheider M, Dragano N. Erhöhtes Risiko eines COVID-19-bedingten Krankenhausaufenthaltes für Arbeitslose: Eine Analyse von Krankenkassendaten von 1,28 Mio. Versicherten in Deutschland [Higher risk of COVID-19 hospitalization for unemployed: an analysis of health insurance data from 1.28 million insured individuals in Germany]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2021 Mar;64(3):314-21. DOI: 10.1007/s00103-021-03280-6

Corresponding author:

Prof. Dr. sc. hum. Ulrike Haug
Leibniz Institute for Prevention Research and
Epidemiology – BIPS, Achterstraße 30, 28359 Bremen,
Germany
haug@leibniz-bips.de

Please cite as

Asendorf M, Reinold J, Schink T, Kollhorst B, Haug U. Abbildbarkeit des sozioökonomischen Status in der pharmakoepidemiologischen Forschungsdatenbank GePaRD: Beschreibung und Anwendung am Beispiel des Zusammenhangs mit Adipositas. *GMS Med Inform Biom Epidemiol*. 2022;18(1):Doc02.
DOI: 10.3205/mibe000235, URN: urn:nbn:de:0183-mibe0002350

This article is freely available from

<https://doi.org/10.3205/mibe000235>

Published: 2022-05-12

Copyright

©2022 Asendorf et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.