SSM - Population Health 11 (2020) 100617

Contents lists available at ScienceDirect

SSM - Population Health

journal homepage: http://www.elsevier.com/locate/ssmph

Review Article

Practice of reporting social characteristics when describing representativeness of epidemiological cohort studies – A rationale for an intersectional perspective

Philipp Jaehn [a,*], Julia Rehling [b], Ronny Klawunn [a,c], Sibille Merz [a], Christine Holmberg [a,d], on behalf of theADVANCE GENDER study group

[a] Brandenburg Medical School Theodor Fontane, Institute of Social Medicine and Epidemiology, Hochstraße 15, 14770, Brandenburgan der Havel, Germany
[b] Umweltbundesamt, Corrensplatz 1, 14195, Berlin, Germany
[c] Hannover Medical School, Institute for Epidemiology, Social Medicine, and Health Systems Research, - OE 5410 -, Carl-Neuberg-Str. 1, 30625, Hannover, Germany
[d] Faculty of Health Sciences, joint Faculty of the Brandenburg University of Technology Cottbus – Senftenberg, the Brandenburg Medical School Theodor Fontane and the University of Potsdam, Germany

ARTICLE INFO

Keywords:
Representativeness
Study participation
Cohort studies
Social location
Intersectionality

ABSTRACT

Representativeness has been defined as the degree of similarity of a study population compared to an external population. To characterize a study population, both health-related and social or demographic features should be considered according to current guidelines. However, little guidance is given on how to describe social complexity of study populations when aiming to conclude on representativeness. We argue that sociological concepts should inform characterizations of study populations in order to increase credibility of conclusions on representativeness. The concept of intersectionality suggests to conceptualize social location as a combination of characteristics such as sex/gender and ethnicity instead of focusing on each feature independently. To contextualize advantages of integrating the concept of intersectionality when investigating representativeness, we reviewed publications that described the baseline population of selected epidemiological cohort studies. Information on the applied methods to characterize the study population was extracted, as well as reported social characteristics. Nearly all reviewed studies reported descriptive statistics of the baseline population and response proportions. In most publications, study populations were characterized according to place of residence, age and sex/gender while other social characteristics were reported irregularly. Differential patterns of representativeness were revealed in analyses that stratified social characteristics by sex/gender or age. Furthermore, the included studies did not explicitly state the theoretical approach that underlay their description of the study population. Intersectionality might be particularly fruitful when applied to descriptions of representativeness, because this concept provides an understanding of social location that has been developed based on situated experiences of people at the intersection of multiple axes of social power relations. An intersectional perspective, hence, contributes to approximate social complexity of study populations and might contribute to increase validity of conclusions on representativeness of population-based studies.

1. Introduction

In A Dictionary of Epidemiology, representativeness is defined as "The degree to which the characteristics of a study (notably, of study subjects and setting, but sometimes also of exposures and outcomes) are similar to those of an external population that did not participate in the study. Representativeness is time-, place-, and context-specific" (Porta, 2014, p. 247). In particular descriptive studies that typically aim to estimate disease burden need to be representative of the target population (Porta, 2014, p. 247; Rothman, Gallacher, & Hatch, 2013).

Current guidelines state that social and demographic characteristics should be used to describe a study population in addition to health-

Abbreviations: NIH, National Institutes of Health; SES, Socio-economic status; USA, United States of America; UK, United Kingdom.

* Corresponding author.
E-mail addresses: philipp.jaehn@mhb-fontane.de (P. Jaehn), Julia.Rehling@uba.de (J. Rehling), Klawunn.Ronny@mh-hannover.de (R. Klawunn), Sibille.Merz@mhb-fontane.de (S. Merz), Christine.Holmberg@mhb-fontane.de (C. Holmberg).

https://doi.org/10.1016/j.ssmph.2020.100617
Received 9 January 2020; Received in revised form 20 May 2020; Accepted 15 June 2020
Available online 20 June 2020

related indicators (International Committee of Medical Journal Editors, 2018; Vandenbroucke et al., 2014). The National Institutes of Health (NIH) issued a policy that obliges investigators of NIH-funded clinical research to report the composition of the study population with regard to sex and ethnic or racial minority groups (National Institutes of Health, 2019). Reviews of studies on substance abuse and physical therapy reported that age and sex are regularly used to describe the study population, while all other social or demographic characteristics are used irregularly (Chevan & Haskvitz, 2015; Gorelick, Montoya, & Johnson, 1998).

To describe representativeness, characteristics of the study population can be compared to the target population using data of the sampling frame, of non-responder-surveys or of censuses (Bethlehem, 2010; Hoffmann et al., 2019). In addition, response proportions can be calculated (Slattery, Edwards, Caan, Kerber, & Potter, 1995), however, high response proportions do not necessarily indicate representativeness, because extended recruitment efforts might increase the relative share of population groups that are most likely to participate (Bethlehem, 2010). On the other hand, the magnitude of bias of an estimate of disease frequency might be lower in studies with high response proportions (Kreienbrock, Pigeot, & Ahrens, 2012).

It has been argued that a sample can be called representative if a sampling technique such as probability sampling has been used (Kruskal & Mosteller, 1979a; 1979b; 1979c). However, despite probability sampling leading to a higher chance of creating a representative sample, study participation is highly likely to distort representativeness (Bergstrand, Vedin, Wilhelmsson, & Wilhelmsen, 1983). Therefore, further investigations of representativeness are needed after recruitment has been completed (Bethlehem, 2010).

If representativeness is understood as a portrayal of a society, we would argue that these investigations should be informed by sociological theory as for example understood by Krieger (Krieger, 2011). Merely stringing together some social or demographic characteristics does not allow an informed analysis of representativeness for complex societies. As a first step to overcome this limitation, we propose to apply an intersectional framework. Intersectionality considers the multiplicity of social location, meaning different intersects of positions individuals act from and are acted on in society (Anthias, 2012; Hankivsky, 2012). Such a framework may be well suited to expand the notion of representativeness of a sample to include a theory-based understanding of the underlying society.

Crenshaw coined the term intersectionality when she illustrated that lived experiences of black women are incomparable to both white females and black males. She supported her argument using case studies of legal discrimination against black women in the US (Crenshaw, 1989). Taking account of situated experiences of multiple discrimination, intercategorical intersectionality posits that characteristics of social location such as race or sex/gender cannot be investigated separately but that combinations of these characteristics need to be the basis of analysis (Bauer, 2014; Bowleg, 2012; Dhamoon & Hankivsky, 2011; McCall, 2005). Intersectionality theory highlights that axes of power relations, which create privileged and oppressed social locations, are interlocking and mutually constituting (Winker & Degele, 2015).

As a corollary, characteristics that are meaningful in an intersectional approach are characteristics related to social differentiation or power (Hankivsky, 2012). The PROGRESS plus framework provides an overview of social groupings for whom anti-discrimination legislation had been enacted in the US (Evans & Brown, 2003; Oliver et al., 2008). Within the framework, 11 groupings have been defined, namely place of residence, ethnicity, occupation, gender, religion, education, socioeconomic status (SES), social capital, disability, age, and sexuality without claiming completeness (Oliver et al., 2008). SES does not refer to a multidimensional construct of income, occupation and education but to material circumstances (Oliver et al., 2008). In this paper, we use sex/gender to express that the embodiment of gender complicates a clear separation of the biological concept of sex and the social concept of gender (Krieger, 2005). Furthermore, we use dis/ability to avoid language that conveys disablism. Finally, it has been suggested that PROGRESS plus might be used to select meaningful social characteristics to describe a study population (Evans & Brown, 2003; O'Neill et al., 2014; Oliver et al., 2008). Characteristics that have shown to be associated with study participation are place of residence, ethnicity, social class, sex, education, income, marital status and age (Bergstrand et al., 1983; Cottler, Zipp, Robins, & Spitznagel, 1987; Neill, Marsden, Matthis, Raspe, & Silman, 1995; Sheikh & Mattingly, 1981; Vernon, Roberts, & Lee, 1984).

To contextualize our theoretical argument, we investigated the practice of reporting social characteristics when describing representativeness at baseline of selected epidemiological cohort studies in a narrative review. We asked if and how study and target population were compared, and which social characteristics of the study population were reported. Finally, we aimed to report results of comparisons of the study populatiom with the target population after stratifying the populations by two or more social characteristics. These strata can be called intersectional strata, because they reveal combinations of social characteristics (Evans, Williams, Onnela, & Subramanian, 2018). By doing so, we intended to explore whether a description of representativeness applying an intersectional perspective might yield results that would have been masked when assuming independence of social characteristics.

2. Methods

2.1. Search strategy

This study was part of the project AdvanceGender (Pöge et al., 2019). We conducted a narrative review of published information about representativeness at baseline of a priori selected epidemiological cohort studies. In a first step, epidemiological cohort studies were included according to two criteria: Studies needed to include a general adult population and baseline recruitment needed to be complete before 2019. In order to analyze descriptions of representativeness of cohort studies, we chose a subjective selection process of individual studies. We considered three sources to search for eligible studies. First, studies that were well-known to the authors were selected. Subsequently, Google and Google Scholar were searched with the terms "cohort study", "cohort profile", and "health" to identify further cohort studies. Finally, websites and methodological publications of already included studies were screened for cross references to other cohort studies. Once identified, the name of the included study, country, number of participants, time of recruitment, and age and sex/gender of participants were recorded.

In the second step, we collected information on the study methodology of each cohort. Therefore, we searched for any kind of methodological publication about representativeness of the identified studies. More precisely, we aimed to include descriptions about study design, study population, study participation, recruitment, response, representativeness or any method of comparison of the study population to the target population. Publications about the baseline population of the cohort studies were included, while publications on attrition or retention during follow-up were excluded. Moreover, any publication that addressed a specific research question was excluded. Only publications in English or German were considered.

To get the information on study methodology we first searched the content of study-websites. Furthermore, we searched publication lists of the study-websites for peer-reviewed publications and grey literature on methodology. Medline, Google, and Google Scholar were additionally searched using the name and acronym of the study and the keywords "cohort profile", "study design", "study description", "study population", "participation", "recruitment", "response" and "representativeness". If none or little information about representativeness at baseline was found, we additionally searched reference lists of publications of the

respective cohort studies for cross references to methodological publications. If two or more publications were found, we included the publication which contained most information. If the reported information concerning one cohort was not congruent, we included all identified sources of information.

2.2. Data extraction

Information on the method of describing representativeness of the study population after recruitment was extracted from the identified websites and publications. Four categories of methods were considered. First, we recorded whether any summary statistic of social characteristics of the study population was presented (A). Reporting response proportions according to Slattery et al. and stratifying them according to age or sex/gender was considered the second method (B) (Slattery et al., 1995). Third, methods using non-responder surveys to compare characteristics of responders to non-responders (C) and, fourth, methods comparing characteristics other than age and sex/gender of the study population to the target population using the sampling frame, census data or data of representative external studies (D) were extracted.

Moreover, any reported social characteristic used to describe the study population was extracted from the identified publications and documented in a spreadsheet. The characteristics as reported in the publications were allocated to one PROGRESS plus group. We compiled a summary for each study indicating PROGRESS plus groups for which at least one characteristic of the study population was reported. We highlighted groups in which one characteristic was stratified by one or more characteristics of another PROGRESS plus group.

Strata of combinations of two or more social characteristics can be called intersectional strata (Evans et al., 2018). Comparisons of intersectional strata of the study population to the target population were summarized and reported following the scheme of PROGRESS plus. An assessment of risk of bias of individual studies was not performed. We address the possibility of publication bias in the Discussion.

Table 1
Characteristics of included cohort studies.

Study name	Acronym	Study design	Country	Restriction to sex/gender	Target age range at baseline	Time of recruitment	Participants at baseline
Framingham Heart Study	FHS	cohort study	United States of America		30–59 years	1948	5209
The National FINRISK Study (1972–1992 surveys)	FINRISK 72-92	cohort study	Finland		25–74 years	1972–1992	45902
The Nurses Health Study I	NHS I	occupational cohort study	United States of America	restriction to females	30–55 years	1976	121700
The Whitehall II Study	Whitehall II	occupational cohort study	United Kingdom		35–55 years	1985–1988	10314
GAZ and Electricité Study	GAZEL	occupational cohort study	France		females: 35–50 years males: 40–50 years	1989	20625
The Avon Longitudinal Study of Parents and Children (mothers cohort)	ALSPAC mothers	birth cohort study	United Kingdom	restriction to females	no age range	1991–1992	13761
European Prospective Investigation into Cancer and Nutrition (international)	EPIC	cohort study	10 European countries		35–70	1992–2000	519978
Västerbotten Intervention Programme (surveys of the years 1992 and 1993)	VIP 92/93	cohort study	Sweden		20–60 years	1992–1993	14188
Monitoring Project on Risk Factors for Chronic Diseases	MORGEN	cohort study	The Netherlands		20–59 years	1993–1997	22769
European Prospective Investigation into Cancer and Nutrition (Germany)	EPIC Germany	cohort study	Germany		women: 35–64 years men: 40–64 years	1994–1998	53162
Cohort of Norway	CONOR	cohort study	Norway		older than 20 years	1994–2003	173236
The National Longitudinal Study of Adolescent to Adult Health	Add Health	cohort study	United States of America		adolescents in high school	1994–1995	90000
Study of Health in Pomerania	SHIP	cohort study	Germany		20–79	1997–2001	4308
Cooperative Health Research in the Region Augsburg (survey of the year 2000)	KORA 2000	cohort study	Germany		25–74 years	2000	4261
Multi-Ethnic Study of Atherosclerosis	MESA	cohort study	United States of America		45–84 years	2000–2002	6814
Heinz Nixdorf Recall Study	HNR	cohort study	Germany		45–75 years	2000–2003	4487
Health, Alcohol and Psychosocial factors in Eastern Europe	HAPIEE	cohort study	Czech Republic, Poland, Russia		45–69 years	2002–2005	36500
China Kadoorie Biobank	CKB	cohort study	China		35–74 years	2004–2008	515681
UK Biobank	UK Biobank	cohort study	United Kingdom		40–69 years	2006–2010	500000
LifeLines Study	LifeLines	cohort study	The Netherlands		older than 6 months	2006–2013	167729
Tromsø 6	Tromsø 6	cohort study	Norway		30–87 years	2007–2008	12984
CARTaGENE Study	CaG	cohort study	Canada		40–69 years	2009–2014	20007
Leben in der Arbeit	LidA	occupational cohort study	Germany		born in 1959 or 1965	2011	6585
CONSTANCES Study	CONSTANCES	occupational cohort study	France		18–69 years	2012–2017	20000

3. Restults

3.1. Included studies and publications

We included 24 epidemiological cohort studies. All included studies ordered chronologically according to their baseline recruitment are displayed in Table 1. The study sizes ranged from 4261 to 519978 participants. The earliest included study started recruiting in 1948. Two studies were restricted to females. 29 publications were identified, among them 26 journal articles and 3 reports from study homepages that met the inclusion criteria (Table 2).

Table 2
Included publications and method of describing representativeness of the study population.

Study	Included publications	Method of describing representativeness of the study population				Response proportion
		A	B	C	D	
FHS	Gordon and Kannel (1968)	X	X	X		69%
FINRISK 72–92	(Borodulin et al., 2017; Harald, Salomaa, Jousilahti, Koskinen, & Vartiainen, 2007)	X	X	X		85%
NHS I	Barton et al. (1980)	X	X			71%
Whitehall II	Marmot et al. (1991)	X	X			73%
GAZEL	Goldberg et al. (2001)	X	X		X	45%
ALSPAC mothers	Fraser et al. (2013)	X	X		X	75%[b]
EPIC	Riboli et al. (2002)	X				not reported
VIP 92/93	Weinehall, Hallgren, Westman, Janlert, and Wall (1998)	X	X		X	57%
MORGEN	Van Loon, Tijhuis, Picavet, Surtees, and Ormel (2003)	X	X	X	X	45%
EPIC Germany	Boeing, Korfmann, and Bergmann (1999)	X	X		X	28%
CONOR	Naess et al. (2008)	X				not reported
Add Health	(Harris, 2013; Kalsbeek, Morris, & Vaughn, 2001)		X			79%[c]
SHIP	(Latza et al., 2004; Völzke et al., 2011)	X	X	X		69%[a]
KORA 2000	(Hoffmann et al., 2004; Rathmann et al., 2003)	X	X	X		65%
MESA	Olson, Bild, Kronmal, and Burke (2016)	X				not reported
HNR	Stang et al. (2005)	X	X	X		53%
HAPIEE	Peasey et al. (2006)	X	X			59%
CKB	Chen et al. (2011)	X				not reported
UK Biobank	Fry et al. (2017)	X	X		X	6%
LifeLines	Klijs et al. (2015)	X			X	not reported
Tromsø 6	Eggen, Mathiesen, Wilsgaard, Jacobsen, and Njolstad (2013)	X	X		X	66%
CaG	Awadalla et al. (2013)	X	X		X	26%[a]
LidA	Hasselhorn et al. (2014)	X	X		X	27%
CONSTANCES	(Zins & Goldberg, 2015; Zins, Goldberg, & team, 2015)	X			X	not reported

Method of describing representativeness of the study population: A) descriptive statistics of the study population; B) response proportions (according to age or sex/gender); C) non-responder survey; D) comparison to the sampling frame, survey or census data (characteristics other than age and sex/gender).

[a] Cooperation proportion.
[b] Proportion of all registered births during study period.
[c] Estimated in in-house survey.

3.2. Method of describing representativeness of the study population

In the identified publications of four cohort studies we could find neither information on response proportions nor any comparison to the target population (Table 2). Four cohort studies exclusively reported aggregated response proportions or response proportions stratified by sex/gender or age. Six cohort studies reported results of a comparison of study participants to participants of a non-responder survey. In eleven studies, comparisons of characteristics of the study population other than age or sex/gender to the target population were conducted using data of the sampling frame, representative censuses or representative survey data. Response proportions ranged from 6% to 84%.

3.3. Reported social characteristics of the study population and allocation to PROGRESS plus groups

The allocation of reported social characteristics to PROGRESS plus groups is summarized in Table 3. All reported characteristics could be allocated to a PROGRESS plus group. The groups “place of residence”, “ethnicity”, “occupation”, “socioeconomic status”, and “social capital” included heterogeneous characteristics.

Fig. 1 shows PROGRESS plus groups for which at least one characteristic has been used to describe the study population and whether these variables had been stratified by variables of other groups. “Sex/gender”, “age” and “place of residence” were reported most frequently. No study reported characteristics of the groups “religion”, “dis/ability” and “sexuality”. If stratification was applied, characteristics were most often stratified by “sex/gender” followed by “age”.

Table 3
Allocation of reported social characteristics to PROGRESS plus groups.

PROGRESS plus group	Reported social characteristics
Place of residence	political regions
	geographical regions
	population size of the region
	population density of the region
	distance to study center
	regional deprivation
Ethnicity	ethnicity
	race
	language
	migration background
	nationality
	place of birth
Occupation	economic sector of occupation
	type of occupation
	unemployment
	marginal employment
	part time employment
	employment grade
	duration of employment
Sex/gender	binary sex/gender
Religion	–
Education	educational level
	duration of education
Socioeconomic status	income
	property ownership
	car ownership
	number of persons living in a room
	type of housing
Social capital	relationship status
	cohabitation
	marital status
	number of children
Age	age-groups
	mean age
	birth cohort
Dis/ability	–
Sexuality	–

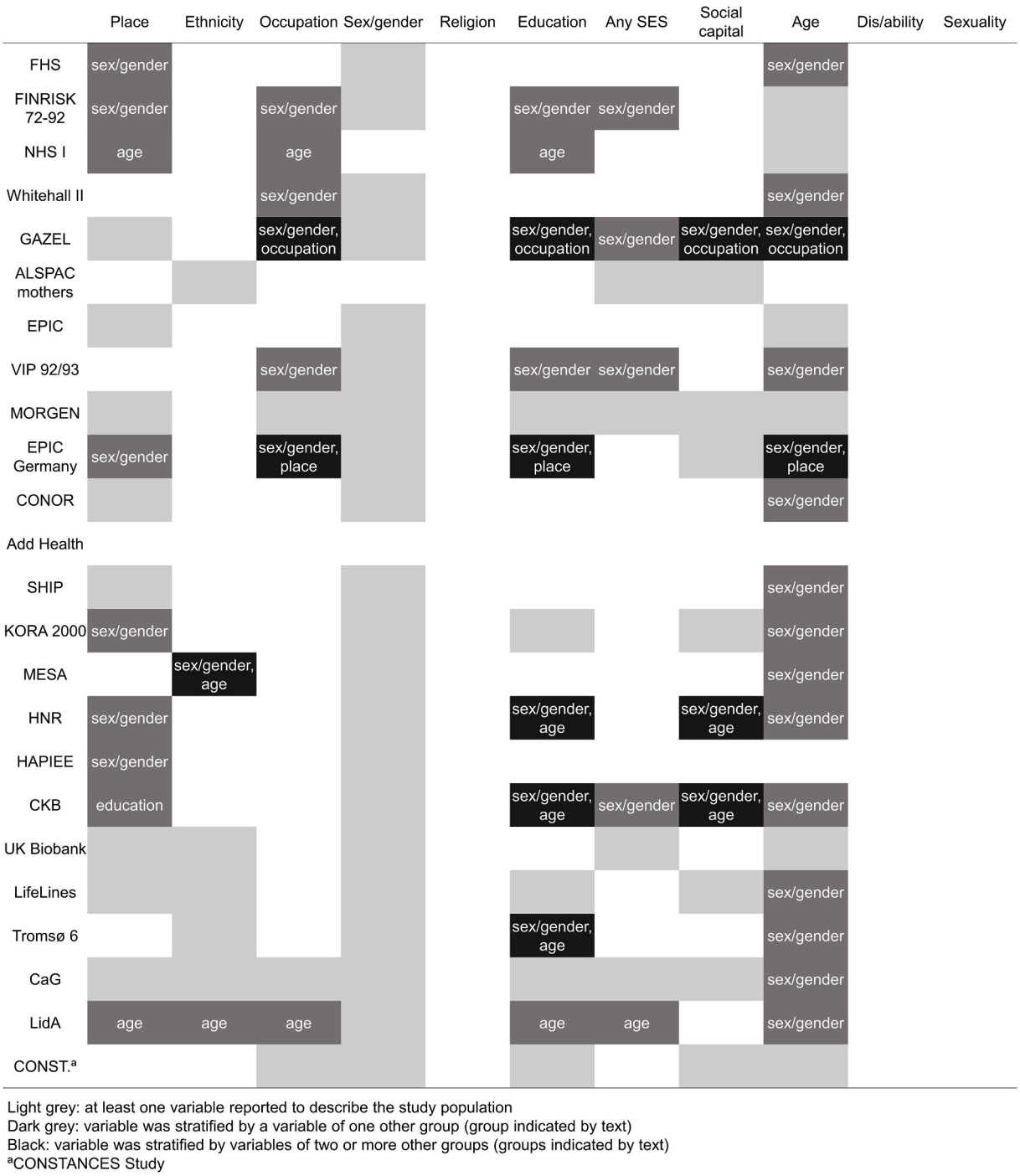

Fig. 1. Summary of PROGRESS plus groups used to describe the study population and their mutual stratification in the included cohort studies.

3.4. Comparisons of intersectional strata of the study population to the target population

3.4.1. Place

Ranges of response rates within the districts of Framingham were similar among women and men in the FHS (Gordon & Kannel, 1968). In the NHS I, overall response proportions in states of the US were similar to response proportions after stratification by age (Barton et al., 1980). EPIC Germany reported a larger female to male ratio of response proportions in Potsdam (city in former East Germany) compared to Heidelberg (city in former West Germany) (Boeing et al., 1999). KORA reported a slightly higher response to a non-responder survey among females in rural areas compared to females in urban areas (Hoffmann et al., 2004). LidA reported that the study population of two birth cohorts was representative for the population in former East and West Germany (Hasselhorn et al., 2014).

3.4.2. Ethnicity

The LidA study stratified nationality by birth cohort and reported a higher representation of foreign nationals in the younger cohort (Hasselhorn et al., 2014).

3.4.3. Occupation

In the NHS I, overall response proportions of the variables field of employment and unemployment status were similar to response proportions after stratification by age (Barton et al., 1980). FINRISK found

that manual workers were underrepresented among women and men (Harald et al., 2007). In GAZEL, employment grade was stratified by sex/gender. Response proportions were approximately double among both female and male managers compared to unskilled workers (Goldberg et al., 2001). Furthermore, time in employment was stratified by sex/gender and employment grade. Only female managers were more likely to participate with increasing length of employment (Goldberg et al., 2001). Conversely, odds of participation decreased only among male unskilled workers with time of employment (Goldberg et al., 2001). In EPIC Germany, the proportion of female blue-collar workers was a quarter compared to the target population, while the proportion of male blue-collar workers was half (Boeing et al., 1999). LidA compared proportions of occupational groups, people in part-time employment and people in marginal work in both birth cohorts to the target population and found no differences (Hasselhorn et al., 2014).

3.4.4. Sex/gender

Stratification by sex/gender has been described in paragraphs about other PROGRESS plus groups.

3.4.5. Religion

Characteristics related to religion were not reported in any study.

3.4.6. Education

Among nurses in the NHS I, overall response proportions according to educational level did not differ from response proportions after stratification by age (Barton et al., 1980). FINRISK reported no differences between study population and target population according to level of education among both women and men (Harald et al., 2007). In GAZEL, education was associated with response among female but not male managers (Goldberg et al., 2001). Among unskilled workers, higher education was associated with higher response among men, but not women (Goldberg et al., 2001). In EPIC Germany and HNR, underrepresentation of people with low educational levels did not differ by sex/gender (Boeing et al., 1999; Stang et al., 2005). LidA reported no differences between study population and target population after stratifying educational levels by birth cohort (Hasselhorn et al., 2014).

3.4.7. SES

Both women and men with a low income were underrepresented in FINRISK (Harald et al., 2007). GAZEL reported that the odds of participation were higher among both males and females living in employer-provided housing compared to those living in private housing (Goldberg et al., 2001). In LidA, income groups stratified by birth cohort were representative of the target population (Hasselhorn et al., 2014).

3.4.8. Social capital

In GAZEL, number of children interacted with sex/gender and employment grade. Only unskilled and skilled male workers were more likely to participate when having more children, while having more children was not associated with participation among male managers or females (Goldberg et al., 2001). In the HNR study, male participants but not female participants were more often married than non-participants (Stang et al., 2005).

3.4.9. Age

In the FHS, men under 45 responded less often than women under 45 (Gordon & Kannel, 1968). GAZEL studied the interaction of age, sex/gender and employment grade (Goldberg et al., 2001). Among men, response decreased with increasing age among unskilled workers only. On the other hand, response decreased with increasing age among female managers (Goldberg et al., 2001). In VIP, the youngest age-group was less represented among both women and men (Weinehall et al., 1998). EPIC Germany stratified response of age-groups by sex/gender and place of residence and found that response among old men and women was higher in Potsdam compared to Heidelberg (Boeing et al., 1999). Stratifying response of age-groups by sex/gender showed that response was lowest among males in the youngest age-groups and females in the oldest age-groups in HNR and SHIP (Latza et al., 2004; Stang et al., 2005). In KORA, young and old men responded less frequently to a non-responder survey compared to women of the same age-group (Hoffmann et al., 2004). The proportion of males of any age-group was reported to be close to the proportions in the target population in Lifelines, while females of younger age-groups were overrepresented (Klijs et al., 2015). Similarly, men of the youngest age-group in CaG showed the lowest response proportion (Awadalla et al., 2013). LidA reported that representativeness of two birth cohorts did not differ between females and males (Hasselhorn et al., 2014). In Tromsø 6, response proportions among women and men were similar across all age-groups with an exception of the oldest age-group. People over 80 years of age in Tromsø 6 showed a lower response among both women and men (Eggen et al., 2013).

3.4.10. Dis/ability

Characteristics related to dis/ability were not reported in any study.

3.4.11. Sexuality

Characteristics related to sexuality were not reported in any study.

4. Discussion

Methods to investigate representativeness of the study population differed between the included cohort studies. Descriptive statistics of the study population and response proportions were most common when describing representativeness. Studies regularly reported sex/gender, age and information on place of residence of the study population which is in line with findings from clinical trials (Chevan & Haskvitz, 2015; Gorelick et al., 1998). Other groups of the PROGRESS plus framework were reported inconsistently. Social characteristics were rarely stratified for each other. If stratification was applied, characteristics were frequently stratified by sex/gender and age. We noted some differential patterns of representativeness after stratifying by a variable of a different PROGRESS plus group. For example, response proportions among women and men were frequently reported to be differential by age. Men responded less often than women in younger age-groups, while differences in response between women and men were less pronounced in older age-groups. Finally, GAZEL reported interaction of employment grade and sex/gender with year of birth, length of employment, educational level and number of children when studying risk factors of non-response (Goldberg et al., 2001).

The identified information of the included cohort studies gives an overview of the practice of reporting on representativeness. Nevertheless, our results summarize the practice of describing representativeness in methodological publications only. We might have missed information on representativeness, if it was published in an article about a specific research question. Additionally, the variability of the identified information on representativeness might be due to heterogeneous study designs. CONOR and EPIC, for example, are large consortia of cohort studies. Results on representativeness of single study centers of CONOR and EPIC are available but were not displayed in the description of the overall project. Add Health, on the other hand, used complex techniques to adjust for non-response while no description of social characteristics of the study population was found in methodological publications (Harris, 2013; Kalsbeek et al., 2001). Furthermore, some characteristics of PROGRESS plus could not be collected due to legal barriers in some studies. For example, the collection of data on ethnicity or religion was banned in several European countries until recently (Simon, 2012). Our findings are not representative of the entire body of epidemiological studies, because the selection of studies was subjective. Moreover, the included studies have been conducted in different time periods and regions. Hence, our results should be understood as a sketch of epidemiological practice across different time periods and social contexts.

Another important aspect that might explain the heterogeneous descriptions of representativeness is that some included studies specifically aimed at obtaining a representative sample, while others focused on low attrition in order to estimate internally valid measures of effect (Collins, 2012). MORGEN, HNR, CaG or Tromsø 6, for example, explicitly aimed at recruiting representative study populations in order to be able to estimate valid measures of disease burden in the target population (Awadalla et al., 2013; Eggen et al., 2013; Stang et al., 2005; Van Loon et al., 2003). For example, MESA, the UK Biobank or CKB were not designed to include a representative study population and instead aimed at studying causal effects (Chen et al., 2011; Fry et al., 2017; Olson et al., 2016). Internally valid effect estimates from unrepresentative cohort studies may be generalizable if effect measure modification of the relationship under study is absent (Rothman et al., 2013). It is important to keep in mind that cohort studies can be designed for these different purposes.

Past analyses of representativeness of social locations are limited, because mostly descriptions of few independent social characteristics were used to characterize the study population and the theoretical rationale of these descriptions has not been stated explicitly. Lisa Bowleg, who was among the first to introduce intersectionality to public health theory, research and policy, illustrated in 2012 that epidemiological or clinical research need to consider tenets of intersectionality when dealing with studies' representation of complex societies (Bowleg, 2012). She noted that the policy of the NIH to include "women and minorities" in all NIH-funded studies was problematic due to the assumed mutual exclusivity of both populations and the multidimensionality and inaccuracy of the term "minorities" (Bowleg, 2012; ; National Institutes of Health, 2019). In the following we present further arguments why intersectionality can advance the practice of investigating representativeness in epidemiological and clinical research.

First, intersectionality posits that intersections of social characteristics constitute unique social locations that cannot be adequately described when analyzing each dimension independently (Bowleg, 2012; Hankivsky, 2012). By making intersectional strata visible through analyses of combinations of multiple social characteristics, complexity of social reality might be represented more adequately (Dhamoon & Hankivsky, 2011; McCall, 2005). Second, intersectionality, calls attention to combinations of characteristics indicating privilege with characteristics indicating disadvantage such as highly educated persons with migration background (Hankivsky, 2012). This perspective is hidden when analyzing single characteristics and shows that almost all strata are combinations of privileged and oppressed categories. By decentering the focus from single characteristics, intersectionality might furthermore contribute to avoid the reinforcement of stereotypes (Dhamoon & Hankivsky, 2011). Consequently, an intersectional perspective highlights that systems of social power are mutually dependent and inseparably related to one another (Dhamoon & Hankivsky, 2011).

This notion of interrelatedness has not been considered in recently developed frameworks that aim to capture dimensions of social location. For example, PROGRESS plus compiled an overview of meaningful characteristics for the description of study populations that moves beyond analyses of commonly studied characteristics such as age, socioeconomic position or sex/gender (Oliver et al., 2008). However, the mutual dependence of these dimensions is not considered. Therefore, frameworks such as PROGRESS plus need to be informed by intersectionality and extended by incorporating perspectives that highlight mutual dependence of all considered categories (Bowleg, 2012; Hankivsky, 2012).

In addition to characteristics investigated in this study, health states like BMI, mental health or HIV status might be considered as further important characteristics in an intersectional approach to representativeness because health might be both a cause of power inequality and a predictor of study participation (O'Hara & Gregg, 2006; Pescosolido & Martin, 2015; Tell et al., 1993; Winker & Degele, 2015). In epidemiological studies, descriptions of health-related characteristics should be part of comprehensive investigations of selection bias (Hernan, Hernandez-Diaz, & Robins, 2004).

Nonetheless, some caveats should be kept in mind when applying an intersectional approach (Bowleg, 2012; Dhamoon & Hankivsky, 2011; McCall, 2005). To start with, combinations of essentialist characteristics are still used and the risk remains that this perspective further emphasizes and shapes difference (McCall, 2005). Moreover, relevant intersections should be identified anew for each social and historical context in which a study is conducted (Dhamoon & Hankivsky, 2011). The use of standard sets of intersections such as a universal application of all PROGRESS plus groups is not recommended (Dhamoon & Hankivsky, 2011). Finally, intersectionality aims at investigating the socio-structural level as well as social processes (Bowleg, 2012), which we neglected to a large part by focusing on social location as an individual characteristic.

An application of intersectionality theory requires the use of novel statistical approaches to analyze multiple interactions. Statistical challenges arise when stratifying a population by several variables because numbers of observations within each stratum decrease geometrically (Kruskal & Mosteller, 1979b). This limitation can be addressed by applications of established and novel statistical methods that aim to operationalize intersectionality. In past studies on intersectional research questions, combinations of two or three variables have been investigated by stratification or by interaction analyses of two to three variables (Mena, Bolte, & Advance Gender Study Group, 2019). Innovative methods that were used in past studies are classification and regression tree analysis or synergy indices (Mena et al., 2019). Recently developed methods include intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA), which can handle small numbers of observations among intersectional strata if the overall sample size is large (Axelsson Fisk et al., 2018;Evans et al., 2018; Hernandez-Yumar et al., 2018; Merlo, 2018). Applications of intersectional MAIHDA to non-responder data, for example, might yield new insights into representativeness across intersectional strata in large population-based studies.

Extended analyses of representativeness are probably most relevant in descriptive epidemiology (Rothman et al., 2013). If differential patterns of response are discovered across intersectional strata, findings might be used to refine calculation of weights to adjust for non-response (Bethlehem, 2010; Kalton & Flores-Cervantes, 2003). Intersectionality might reveal an underrepresentation of multiply marginalised groups, which needs to be considered when interpreting research results. This knowledge is crucial, as marginalised groups have been historically excluded from research and often suffer from a high burden of disease (Larson, 1994). Therefore, findings might be used for the development of targeted recruitment strategies. However, specifically targeting vulnerable social groups might reinforce social separation and should only be considered after careful evaluation (Epstein, 2008).

Knowledge on differential response patterns might also be useful to evaluate risk for selection bias of measures of effect and population impact. Large longitudinal studies with low representativeness might be at higher risk to yield estimates of effect that might not be generalizable, especially when effect measure modification is present (Stang & Jöckel, 2014). It has been shown that measures of effect might be biased if both effect measure modification and differential response according to the modifying factor are present (Jöckel & Stang, 2013). Effect measure modification might be anticipated for example when investigating research questions in social epidemiology (Patil, Porche, Shippen, Dallenbach, & Fortuna, 2018; Rothman et al., 2013). While drawing generalizable conclusions on causal relationships is possible in studies with low representativeness, measures of population impact need to be estimated in representative studies (Rothman et al., 2013). Applying an intersectional analysis of representativeness could add an additional layer of certainty when aiming to draw generalizable conclusions on causal effects and population impact of risk factors.

Finally, epidemiology has gradually moved past estimating

population averages towards mapping heterogeneity of disease burden in a population as well as studying heterogeneity of effects (Merlo, 2014, 2018). MAIHDA, for example, aids to map heterogeneity of disease frequencies among intersectional strata of a population (Evans et al., 2018; Persmark et al., 2019). If epidemiological studies, on the other hand, aim at estimating effect heterogeneity across population subgroups, study designs might need to adapt. It has been suggested, for example, to study several homogeneous population strata that are heterogeneous in relation to each other rather than representative samples of the general population to estimate effect heterogeneity (Merlo, Mulinari, Wemrell, Subramanian, & Hedblad, 2017).

In conclusion, applying the concept of intersectionality might approximate social complexity when investigating representativeness by taking account of the multiplicity of social location. An intensified integration of sociological concepts might advance our understanding of representativeness and current epidemiological practice.

Author contributions

PJ conceptualized the research question, curated data, analyzed the data and drafted the manuscript. JR and RK were involved in data curation and reviewed and edited the manuscript. SM reviewed and edited the manuscript. CH acquired funding, administered the project, conceptualized the research question and reviewed and edited the manuscript.

Funding

This work was supported by the German Federal Ministry of Education and Research [reference no. 01GL1710a].

Ethics approval

The study was approved by the ethics board of the Brandenburg Medical School (ref. E−01-20180529).

Declaration of competing interest

None.

Acknowledgements

We want to thank the staff of the included studies and all members of the joint project AdvanceGender for their support. Members of the AdvanceGender study group: University of Bremen, Institute of Public Health and Nursing Research, Department of Social Epidemiology (Gabriele Bolte, Emily Mena), Robert Koch Institute (Alexander Rommel, Anke-Christine Saß, Kathleen Pöge, Sarah Strasser), Brandenburg Medical School Theodor Fontane, Institute of Social Medicine and Epidemiology (Christine Holmberg, Sibille Merz, Philipp Jaehn).

References

Anthias, F. (2012). Hierarchies of social location, class and intersectionality: Towards a translocational frame. *International Sociology, 28*(1), 121–138. https://doi.org/10.1177/0268580912463155.

Awadalla, P., Boileau, C., Payette, Y., Idaghdour, Y., Goulet, J. P., Knoppers, B., … Project, C. A. (2013). Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. *International Journal of Epidemiology, 42*(5), 1285–1299. https://doi.org/10.1093/ije/dys160.

Axelsson Fisk, S., Mulinari, S., Wemrell, M., Leckie, G., Perez Vicente, R., & Merlo, J. (2018). Chronic Obstructive Pulmonary Disease in Sweden: An intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy. *SSM Popul Health, 4*, 334–346. https://doi.org/10.1016/j.ssmph.2018.03.005.

Barton, J., Bain, C., Hennekens, C. H., Rosner, B., Belanger, C., Roth, A., et al. (1980). Characteristics of respondents and non-respondents to a mailed questionnaire. *American Journal of Public Health, 70*(No. 8), 823–825.

Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity. *Social Science & Medicine, 110*, 10–17. https://doi.org/10.1016/j.socscimed.2014.03.022.

Bergstrand, R., Vedin, A., Wilhelmsson, C., & Wilhelmsen, L. (1983). Bias due to non-participation and heterogenous sub-groups in population surveys. *Journal of Chronic Diseases, 36*(10), 725–728.

Bethlehem, J. (2010). *New developments in survey data collection methodology for official statistics (Discussion paper 10010). The Hague/Heerlen: Statistics Netherlands*.

Boeing, H., Korfmann, A., & Bergmann, M. M. (1999). Recruitment procedures of EPIC-Germany. *Annals of Nutrition & Metabolism, 43*, 205–215.

Borodulin, K., Tolonen, H., Jousilahti, P., Jula, A., Juolevi, A., Koskinen, S., … Vartiainen, E. (2017). Cohort profile: The national FINRISK study. *International Journal of Epidemiology, 47*(3). https://doi.org/10.1093/ije/dyx239, 696-696i.

Bowleg, L. (2012). The problem with the phrase women and minorities: Intersectionality-an important theoretical framework for public health. *American Journal of Public Health, 102*(7), 1267–1273. https://doi.org/10.2105/ajph.2012.300750.

Chen, Z., Chen, J., Collins, R., Guo, Y., Peto, R., Wu, F., … China Kadoorie Biobank collaborative, g.. (2011). China kadoorie biobank of 0.5 million people: Survey methods, baseline characteristics and long-term follow-up. *International Journal of Epidemiology, 40*(6), 1652–1666. https://doi.org/10.1093/ije/dyr120.

Chevan, J., & Haskvitz, E. M. (2015). Reported characteristics of participants in physical therapy-related clinical trials. *Physical Therapy, 95*(6), 884–890. https://doi.org/10.2522/ptj.20140256.

Collins, R. (2012). What makes UK Biobank special? *Lancet, 379*(9822), 1173–1174.

Cottler, L. B., Zipp, J. F., Robins, L. N., & Spitznagel, E. L. (1987). Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey. *American Journal of Epidemiology, 125*(2), 329–339. https://doi.org/10.1093/oxfordjournals.aje.a114534.

Crenshaw, K. (1989). *Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics* (Vol. 1989, pp. 139–167). Univ Chic Leg Forum. Article 8.

Dhamoon, R. K., & Hankivsky, O. (2011). Why the theory and practice of intersectionality matter to health reasearch and policy. In *Health inequities in Canada : Intersectional frameworks and practices*. Vancouver: UBC Press.

Eggen, A. E., Mathiesen, E. B., Wilsgaard, T., Jacobsen, B. K., & Njolstad, I. (2013). The sixth survey of the tromso study (tromso 6) in 2007-08: Collaborative research in the interface between clinical medicine and epidemiology: Study objectives, design, data collection procedures, and attendance in a multipurpose population-based health survey. *Scandinavian Journal of Public Health, 41*(1), 65–80. https://doi.org/10.1177/1403494812469851.

Epstein, S. (2008). The rise of 'recruitmentology': Clinical research, racial knowledge, and the politics of inclusion and difference. *Social Studies of Science, 38*(5), 801–832. https://doi.org/10.1177/0306312708091930.

Evans, T., & Brown, H. (2003). Road traffic crashes: Operationalizing equity in the context of health sector reform. *Injury Control and Safety Promotion, 10*(1-2), 11–12. https://doi.org/10.1076/icsp.10.1.11.14117.

Evans, C. R., Williams, D. R., Onnela, J.-P., & Subramanian, S. V. (2018). A multilevel approach to modeling health inequalities at the intersection of multiple social identities. *Social Science & Medicine, 203*, 64–73. https://doi.org/10.1016/j.socscimed.2017.11.011.

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., … Lawlor, D. A. (2013). Cohort profile: The avon longitudinal study of parents and children: ALSPAC mothers cohort. *International Journal of Epidemiology, 42*(1), 97–110. https://doi.org/10.1093/ije/dys066.

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., … Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *American Journal of Epidemiology, 186*(9), 1026–1034. https://doi.org/10.1093/aje/kwx246.

Goldberg, M., Chastang, J. F., Leclerc, A., Zins, M., Bonenfant, S., Bugel, I., … Imbernon, E. (2001). Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: A prospective study of the French GAZEL cohort and its target population. *American Journal of Epidemiology, 154*(4), 373–384.

Gordon, R., & Kannel, W. B. (1968). *The frameingham study. An epidemiological investigation of cardiovascular disease*. Retrieved from https://biolincc.nhlbi.nih.gov/media/studies/framcohort/An_Epidemiological_Investigation_of_Cardiovascular_Disease_(Sections_1-2).pdf accessed 22.01.2019.

Gorelick, D. A., Montoya, I. D., & Johnson, E. O. (1998). Sociodemographic representation in published studies of cocaine abuse pharmacotherapy. *Drug and Alcohol Dependence, 49*(2), 89–93.

Hankivsky, O. (2012). Women's health, men's health, and gender and health: Implications of intersectionality. *Social Science & Medicine, 74*(11), 1712–1720. https://doi.org/10.1016/j.socscimed.2011.11.029.

Harald, K., Salomaa, V., Jousilahti, P., Koskinen, S., & Vartiainen, E. (2007). Non-participation and mortality in different socioeconomic groups: The FINRISK population surveys in 1972-92. *Journal of Epidemiology & Community Health, 61*(5), 449–454. https://doi.org/10.1136/jech.2006.049908.

Harris, K. (2013). *The add health study: Design and accomplishments*. https://doi.org/10.17615/C6TW87. Retrieved from, accessed 23.04.2019.

Hasselhorn, H. M., Peter, R., Rauch, A., Schroder, H., Swart, E., Bender, S., … Muller, B. H. (2014). Cohort profile: The lidA cohort study-a German cohort study on work, age, health and work participation. *International Journal of Epidemiology, 43*(6), 1736–1749. https://doi.org/10.1093/ije/dyu021.

Hernandez-Yumar, A., Wemrell, M., Abasolo Alesson, I., Gonzalez Lopez-Valcarcel, B., Leckie, G., & Merlo, J. (2018). Socioeconomic differences in body mass index in Spain: An intersectional multilevel analysis of individual heterogeneity and

discriminatory accuracy. *PloS One, 13*(12), Article e0208624. https://doi.org/10.1371/journal.pone.0208624.
Hernan, M. A., Hernandez-Diaz, S., & Robins, J. M. (2004). A structural approach to selection bias. *Epidemiology, 15*(5), 615–625.
Hoffmann, W., Latza, U., Baumeister, S. E., Brünger, M., Buttmann-Schweiger, N., Hardt, J., et al. (2019). Guidelines and recommendations for ensuring good epidemiological practice (GEP): A guideline developed by the German society for epidemiology. *European Journal of Epidemiology, 34*(3), 301–317. https://doi.org/10.1007/s10654-019-00500-x.
Hoffmann, W., Terschuren, C., Holle, R., Kamtsiuris, P., Bergmann, M., Kroke, A., … Latza, U. (2004). Zum problem der Response in epidemiologischen studien in deutschland (teil II). *Das Gesundheitswesen, 66*, 482–491. https://doi.org/10.1055/s-2004-813094.
International Committee of Medical Journal. (2018). In *Recommendations for the conduct, reporting, editing and publication of scholarly work in medical journals*. Retrieved from http://www.ICMJE.org accessed 03.06.2019.
Jöckel, K. H., & Stang, A. (2013). Cohort studies with low baseline response may not be generalisable to populations with different exposure distributions. *European Journal of Epidemiology, 28*(3), 223–227. https://doi.org/10.1007/s10654-013-9782-2.
Kalsbeek, W. D., Morris, C. B., & Vaughn, B. J. (2001). Effects of nonresponse on the mean squared error of estimates from a longitudinal study. In *Paper presented at the annual meeting of the American statistical assocation*. Joint Statistical Meetings.
Kalton, G., & Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics, 19*(2), 81.
Klijs, B., Scholtens, S., Mandemakers, J. J., Snieder, H., Stolk, R. P., & Smidt, N. (2015). Representativeness of the LifeLines cohort study. *PloS One, 10*(9), Article e0137203. https://doi.org/10.1371/journal.pone.0137203.
Kreienbrock, L., Pigeot, I., & Ahrens, W. (2012). *Epidemiologische methoden*. Berlin, Heidelberg: Springer-Verlag [in German].
Krieger, N. (2005). Embodiment: A conceptual glossary for epidemiology. *Journal of Epidemiology & Community Health, 59*(5), 350–355. https://doi.org/10.1136/jech.2004.024562.
Krieger, N. (2011). *Epidemiology and the people's health: Theory and context*. New York, NY: Oxford University Press.
Kruskal, W., & Mosteller, F. (1979a). Representative sampling, I: Non-scientific literature. *International Statistical Review, 47*(1), 13–24. https://doi.org/10.2307/1403202.
Kruskal, W., & Mosteller, F. (1979b). Representative sampling, II: Scientific literature, excluding statistics. *International Statistical Review, 47*(2), 111–127. https://doi.org/10.2307/1402564.
Kruskal, W., & Mosteller, F. (1979c). Representative sampling, III: The current statistical literature. *International Statistical Review, 47*(3), 245–265. https://doi.org/10.2307/1402647.
Larson, E. (1994). Exclusion of certain groups from clinical research. *Image - the Journal of Nursing Scholarship, 26*(3), 185–190. https://doi.org/10.1111/j.1547-5069.1994.tb00311.x.
Latza, U., Stang, A., Bergmann, M., Kroke, A., Sauer, S., Holle, R., … Hoffmann, W. (2004). Zum problem der Response in epidemiologischen studien in deutschland (teil I). *Das Gesundheitswesen, 66*(5), 326–336. https://doi.org/10.1055/s-2004-813093.
Marmot, M., Smith, G. D., Stansfeld, S., Patel, C., North, F., Head, J., … Feeney, A. (1991). Health inequities among British civil servants: The whitehall II study. *The Lancet, 337*(8), 1387–1393.
McCall, L. (2005). The complexity of intersectionality. *Signs, 30*(3), 1772–1800.
Mena, E., Bolte, G., & Advance Gender Study Group. (2019). Intersectionality-based quantitative health research and sex/gender sensitivity: A scoping review. *International Journal for Equity in Health, 18*(1), 199. https://doi.org/10.1186/s12939-019-1098-8.
Merlo, J. (2014). Invited commentary: Multilevel analysis of individual heterogeneity-a fundamental critique of the current probabilistic risk factor epidemiology. *American Journal of Epidemiology, 180*(2), 208–212. https://doi.org/10.1093/aje/kwu108. discussion 213-204.
Merlo, J. (2018). Multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) within an intersectional framework. *Social Science & Medicine, 203*, 74–80. https://doi.org/10.1016/j.socscimed.2017.12.026.
Merlo, J., Mulinari, S., Wemrell, M., Subramanian, S. V., & Hedblad, B. (2017). The tyranny of the averages and the indiscriminate use of risk factors in public health: The case of coronary heart disease. *SSM Popul Health, 3*, 684–698. https://doi.org/10.1016/j.ssmph.2017.08.005.
Naess, O., Sogaard, A. J., Arnesen, E., Beckstrom, A. C., Bjertness, E., Engeland, A., … Aamodt, G. (2008). Cohort profile: Cohort of Norway (CONOR). *International Journal of Epidemiology, 37*(3), 481–485. https://doi.org/10.1093/ije/dym217.
National Institutes of Health. (2019). *NIH policy and guidelines on the inclusion of women and minorities as subjects in clinical research*. Retrieved from www.grants.nih.gov/grants/funding/women_min/guidelines.htm accessed 15.01.2019.
Neill, T. W., Marsden, D., Matthis, C., Raspe, H., & Silman, A. J. (1995). Survey response rates: National and regional differences in a European multicentre study of vertebral osteoporosis. *Journal of Epidemiology & Community Health, 49*(1), 87. https://doi.org/10.1136/jech.49.1.87.
O'Hara, L., & Gregg, J. (2006). The war on obesity: A social determinant of health. *Health Promotion Journal of Australia, 17*(3), 260–263.
O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., et al. (2014). Applying an equity lens to interventions: Using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology, 67*(1), 56–64. https://doi.org/10.1016/j.jclinepi.2013.08.005.
Oliver, S., Kavanagh, J., Caird, J., Lorenc, T., Oliver, K., & Harden, A. (2008). *Health promotion, inequalities and young people's health (EPPI-Centre report no. 1611)*. Retrieved from https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2410 accessed 06.02.2019.
Olson, J. L., Bild, D. E., Kronmal, R. A., & Burke, G. L. (2016). Legacy of MESA. *Glob Heart, 11*(3), 269–274. https://doi.org/10.1016/j.gheart.2016.08.004.
Patil, P. A., Porche, M. V., Shippen, N. A., Dallenbach, N. T., & Fortuna, L. R. (2018). Which girls, which boys? The intersectional risk for depression by race and ethnicity, and gender in the U.S. *Clinical Psychology Review, 66*, 51–68.
Peasey, A., Bobak, M., Kubinova, R., Malyutina, S., Pajak, A., Tamosiunas, A., … Marmot, M. (2006). Determinants of cardiovascular disease and other non-communicable diseases in central and eastern europe: Rationale and design of the HAPIEE study. *BMC Public Health, 6*, 255. https://doi.org/10.1186/1471-2458-6-255.
Persmark, A., Wemrell, M., Zettermark, S., Leckie, G., Subramanian, S. V., & Merlo, J. (2019). Precision public health: Mapping socioeconomic disparities in opioid dispensations at Swedish pharmacies by multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). *PloS One, 14*(8), Article e0220322. https://doi.org/10.1371/journal.pone.0220322.
Pescosolido, B. A., & Martin, J. K. (2015). The stigma complex. *Annual Review of Sociology, 41*, 87–116. https://doi.org/10.1146/annurev-soc-071312-145702.
Pöge, K., Rommel, A., Mena, E., Holmberg, C., Sass, A. C., & Bolte, G. (2019). [AdvanceGender-Joint project for sex/gender-sensitive and intersectional research and health reporting]. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz, 62*(1), 102–107. https://doi.org/10.1007/s00103-018-2855-3.
Porta, M. (2014). *A dictionary of epidemiology*. New York, NY: Oxford university press.
Rathmann, W., Haastert, B., Icks, A., Loewel, H., Meisinger, C., Holle, R., et al. (2003). High prevalence of undiagnosed diabetes mellitus in Southern Germany: Target population for efficient screening. *The KORA survey 2000. Diabetologia, 46*, 182–189. https://doi.org/10.1007/s00125-002-1025-0.
Riboli, E., Hunt, K. J., Slimani, N., Ferrari, P., Norat, T., Fahey, M., … Saracci, R. (2002). European prospective investigation into cancer and nutrition (EPIC): Study populations and data collection. *Public Health Nutrition, 5*(6B), 1113–1124. https://doi.org/10.1079/PHN2002394.
Rothman, K. J., Gallacher, J. E., & Hatch, E. E. (2013). Why representativeness should be avoided. *International Journal of Epidemiology, 42*(4), 1012–1014. https://doi.org/10.1093/ije/dys223.
Sheikh, K., & Mattingly, S. (1981). Investigating non-response bias in mail surveys. *Journal of Epidemiology & Community Health, 35*(4), 293–296. https://doi.org/10.1136/jech.35.4.293.
Simon, P. (2012). Collecting ethnic statistics in europe: A review. *Ethnic and Racial Studies, 35*, 1366–1391. https://doi.org/10.1080/01419870.2011.607507.
Slattery, M. L., Edwards, S. L., Caan, B. J., Kerber, R., & Potter, J. D. (1995). Response rates among control subjects in case-control studies. *Annals of Epidemiology, 5*, 245–249.
Stang, A., & Jöckel, K. H. (2014). Avoidance of representativeness in presence of effect modification. *International Journal of Epidemiology, 43*(2), 630–631. https://doi.org/10.1093/ije/dyt263.
Stang, A., Moebus, S., Dragano, N., Beck, E. M., Möhlenkamp, S., Schmermund, A., … Group, H. N. R. S. I.. (2005). Baseline recruitment and analyses of nonresponse of the Heinz Nixdorf Recall Study: Identifiability of phone numbers as the major determinant of response. *European Journal of Epidemiology, 20*, 489–496. https://doi.org/10.1007/s10654-005-5529-z.
Tell, G. S., Fried, L. P., Hermanson, B., Manolio, T. A., Newman, A. B., & Borhani, N. O. (1993). Recruitment of adults 65 years and older as participants in the cardiovascular health study. *Annals of Epidemiology, 3*(4), 358–366.
Van Loon, A. J., Tijhuis, M., Picavet, H. S., Surtees, P. G., & Ormel, J. (2003). Survey non-response in The Netherlands: Effects on prevalence estimates and associations. *Annals of Epidemiology, 13*(2), 105–110. https://doi.org/10.1016/s1047-2797(02)00257-0.
Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gotzsche, P. C., Mulrow, C. D., Pocock, S. J., … Egger, M. (2014). Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. *International Journal of Surgery, 12*(12), 1500–1524. https://doi.org/10.1016/j.ijsu.2014.07.014.
Vernon, S. W., Roberts, R. E., & Lee, E. S. (1984). Ethnic status and participation in longitudinal health surveys. *American Journal of Epidemiology, 119*(1), 99–113. https://doi.org/10.1093/oxfordjournals.aje.a113731.
Völzke, H., Alte, D., Schmidt, C. O., Radke, D., Lorbeer, R., Friedrich, N., … Hoffmann, W. (2011). Cohort profile: The study of health in pomerania. *International Journal of Epidemiology, 40*(2), 294–307. https://doi.org/10.1093/ije/dyp394.
Weinehall, L., Hallgren, C. G., Westman, G., Janlert, U., & Wall, S. (1998). Reduction of selection bias in primary prevention of cardiovascular disease through involvement of primary health care. *Scandinavian Journal of Primary Health Care, 16*(3), 171–176. https://doi.org/10.1080/028134398750003133.
Winker, G., & Degele, N. (2015). *Intersektionalität: Zur analyse sozialer ungleichheiten*. Bielefeld: transcript Verlag [in German].
Zins, M., & Goldberg, M. (2015). *The CONSTANCES cohort*. Retrieved from http://www.constances.fr/assets/pdf/Scientific-protocol-01-2015.pdf accessed 22.01.2019.
Zins, M., Goldberg, M., & team, C. (2015). The French CONSTANCES population-based cohort: Design, inclusion and follow-up. *European Journal of Epidemiology, 30*(12), 1317–1328. https://doi.org/10.1007/s10654-015-0096-4.